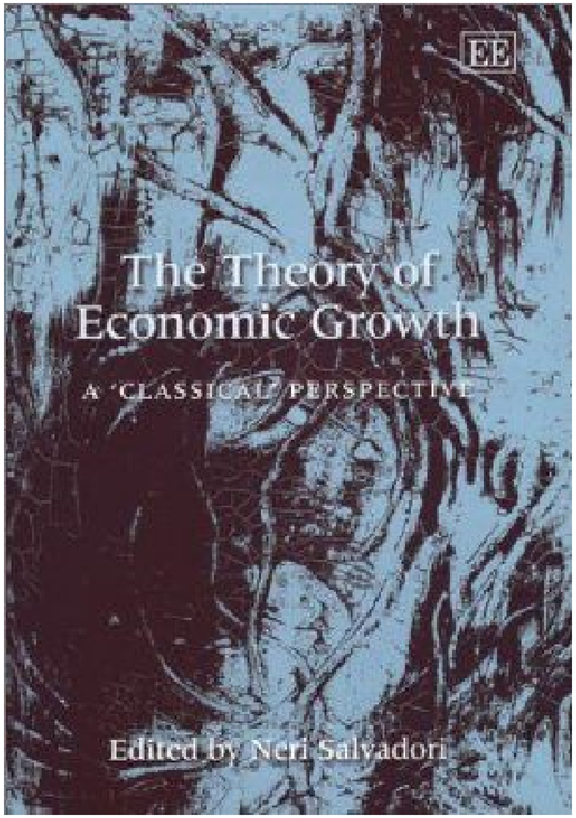


The Theory of Economic Growth: a 'Classical' Perspective

Edited by

Neri Salvadori

University of Pisa, Italy

Edward Elgar
Cheltenham, UK • Northampton, MA, USA

Contents

List of contributors

Mauro Caminati (University of Siena, Italy)

Salvatore Capasso (University of Napoli "Federico II", Italy)

Maria Rosaria Carillo (University of Napoli "Parthenope", Italy)

Pasquale Commendatore (University of Napoli "Federico II", Italy)

Salvatore D'Acunto (University of Napoli "Federico II", Italy)

Antonio D'Agata (University of Catania, Italy)

Luciano Fanti (University of Pisa, Italy)

Davide Fiaschi (University of Pisa, Italy)

Giuseppe Freni (University of Napoli "Parthenope", Italy)

Maria Daniela Giammanco (University of Catania, Italy)

Fausto Gozzi (University of Rome "La Sapienza", Italy)

Heinz D. Kurz (University of Graz, Autriche)

Andrea Lavezzi (University of Pisa, Italy)

Tommaso Luzzati (University of Pisa, Italy)

Piero Manfredi (University of Pisa, Italy)

Carlo Panico (University of Napoli "Federico II", Italy)

Fabio Petri (University of Siena, Italy)

Antonio Pinto (University of Napoli "Federico II", Italy)

Mario Pomini (University of Verona, Italy)

Neri Salvadori (University of Pisa)

Grazia D. Santangelo (University of Catania, Italy)

Rodolfo Signorino (University of Napoli "Federico II")

Serena Sordi (University of Siena, Italy)

Introduction

Neri Salvadori

Interest in the study of economic growth has experienced remarkable ups and downs in the history of economics. It was central in Classical political economy from Adam Smith to David Ricardo, and then in its 'critique' by Karl Marx, but moved to the periphery during the so-called 'marginal revolution'. John von Neumann's growth model and Roy Harrod's attempt to generalise Keynes's principle of effective demand to the long run re-ignited interest in growth theory. Following the publication of papers by Robert Solow and Nicholas Kaldor in the mid 1950s, growth theory became one of the central topics of the economics profession until the early 1970s. After a decade of dormancy, since the mid 1980s, economic growth has once again become a central topic in economic theorising. The recent theory is called 'endogenous growth theory', since according to it the growth rate is determined from within the model and is not given as an exogenous variable.

This book is the main product of a research group on the theory of growth and the relation between modern growth theory and 'Classical' growth theory. The scholars involved were motivated to this task not only by the emergence at the end of the 1980s and the rapid development of the literature on economic growth, but also by the contributions of Kurz and Salvadori (1998b, 1999) who have shown that the logical structure underlying most of the early models of endogenous growth is very similar to the logical structure of 'Classical' growth models. Put schematically, in the latter a given real wage rate determines (together with the technological data) the rate of profits and thus, through the saving-investment mechanism, the rate of growth; in the modern literature, 'human capital' or 'knowledge' works in the same way since there is a 'technology' producing them, exactly like the real wage rate 'produced' labour in the analyses of the Classical economists. The scholars involved have also investigated the connection between the Classical economists and the modern theories of growth in the analysis of competition, technical change, economic cycles, and financial intermediation.

The readers may ask themselves whether classifying economic ideas in distinct analytical approaches to certain economic problems and even in different schools of economic thought is a futile enterprise. The title of this book implies that its authors think that it is not. We rather hold the view that there is a theory that may, for good reasons, be called 'Classical' economics as distinct from other kinds of economics, in particular 'Neoclassical' economics and 'Keynesian' economics. This view could immediately be challenged with the indisputable heterogeneity and multi-layeredness of the writings of authors in these groups. Moreover, whilst regarding some aspects an author might be classified in one group, regarding some other aspects he or she might be classified in another group. Therefore, I wish to make it clear from the outset that we are not so much concerned with elaborating a classification of authors, which in some cases would be an extremely difficult, if not impossible task. We are rather concerned with classifying various analytical approaches to dealing with certain economic problems. Our interest in these approaches is not dominantly historical; we rather consider them as containing the key to a better explanation of important economic phenomena. Our concern with classical economics is therefore primarily a concern with its analytical potential which in our view has not yet been fully explored.

The book opens with a chapter by Kurz and Salvadori that summarises their previous contributions and clarifies what we mean by 'Classical' and 'Neoclassical' economics. Chapters 2 and 3 complete this methodological analysis. Antonio D'Agata and Giuseppe Freni insert also 'Keynesian' economics into the picture and find some other connections among these schools of thought. Mario Pomini studies the emergence of endogenous growth theory (as opposed to Neoclassical growth theory) from the point of view of Lakatosian categories. These chapters isolate and compare the logical structures and the methodological underpinnings of old and new growth theories. They provide some well-defined guidelines that address the analysis developed in the following chapters.

Chapters 4–9 analyse in greater detail the above-mentioned schools of thought: Classical, Keynesian, Neoclassical. Chapter 10, by Santangelo, surveys the evolutionary point of view on growth and thus complements Chapters 4–9. Chapter 4, by Giuseppe Freni, Fausto Gozzi, and Neri Salvadori, can be read as an analysis of the problems that the extension to a multi-sector economy poses for endogenous growth theorists, but it can also be read both as a restatement of some solutions proposed by the theory of production of 'Classical' orientation (see Kurz and Salvadori, 1995) and as a complement to this theory when the growth rate is negative and depreciation is by evaporation. Chapter 5, by Davide Fiaschi and Rodolfo Signorino, investigates a problem concerning the 'Classical' growth model that has rarely been on the agenda of scholars interested in modern developments of the 'Classical' school (but see Pasinetti, 1981, pp. 69–70; 1993): the problem of consumption patterns. Chapter 6 is a broad survey on 'Keynesian' theories

of growth. Pasquale Commendatore, Salvatore D'Acunto, Carlo Panico, and Antonio Pinto have gone to great lengths to produce a comprehensive analysis of all the literature on the issue. Chapter 7 on Say's law, by Fabio Petri, complements this analysis. As is argued in the first chapter of this book, the fact that the endowments of all resources, including capital and labour, are among the data of neoclassical theory imposes that this theory can consider growth only as exogenously directed. However, a sort of alternative exists; it consists in complementing neoclassical theory with a theory modelling the evolution of some endowments. Chapters 8 and 9 perform this task. Piero Manfredi and Luciano Fanti provide an analysis of the dynamics of the working population within the Solovian model. Maria Rosaria Carillo studies the changes in the efficiency of work connected with social factors, as opposed to economic factors.

Thus Chapters 3–10 are mainly devoted to a 'vertical' or in-depth analysis of four schools of thought, the 'Classical', the 'Keynesian', the 'Neoclassical' and the 'Evolutionary' School. By contrast, the remaining chapters of the book are devoted to a 'horizontal' analysis of a number of items connected with growth. Chapter 11, by Antonio D'Agata, explores the problem of legal barriers to entry and rent-seeking in Smith and in the modern theory of growth. Chapters 12 and 13 investigate the problem of technical change: Mauro Caminati proposes an ingenious method to classify the modern literature whereas Maria Daniela Giammanco compares recent results with some features that characterise the analysis of technical change proposed by Marx. Chapter 14, by Andrea Mario Lavezzi, compares the modern contributions on the division of labour with the old literature, mainly Adam Smith and Allyn Young. Chapters 15 and 16 analyse the connection between growth and cycles: Serena Sordi surveys the macrodynamic models whereas Davide Fiaschi and Serena Sordi survey the more recent literature on this topic. Tommaso Luzzati, in Chapter 17, is concerned with the questions that the environment poses for growth theorists. Finally, Chapter 18, by Salvatore Capasso, investigates the problems connected with the existence of financial intermediation.

1. Theories of economic growth: old and new[*]

Heinz D. Kurz and Neri Salvadori

1.1. INTRODUCTION

Ever since the inception of systematic economic analysis at the time of the classical economists from William Petty to David Ricardo the problem of economic growth – its sources, forms and effects – was high on the agenda of economists. In the real world the problem and the fact of economic growth is, of course, of much longer standing. Even in the more or less stationary economies of antiquity the possibility, if not the fact, of economic expansion lingers at the back of certain considerations. Clay tablets from Mesopotamia provide information about social productivity by means of a simple input–output calculation in terms of barley. The main question concerned the *surplus product* of barley the ancient society was able to generate, that is, the excess of total output in a year with a normal harvest over the amount of input of barley as seed or as a means of subsistence for labourers plus any other inputs needed in the society measured in terms of barley. From the Surplus Rate, that is, the ratio of Surplus Product to Necessary Input, it is obviously only a small step intellectually, but a huge step historically, to the concept of the *rate of growth*. This step was taken, at the latest, by economists in the seventeenth century, most notably William Petty.

This chapter is devoted to a brief discussion of the characteristic features of a selection of contributions to the problem under consideration. It summarizes previous contributions by the same authors. The interested reader can see more detailed analyses in Kurz and Salvadori (1998b, 1999). Section 1.2 summarizes some crucial features of Adam Smith's views on capital accumulation and economic growth. The emphasis is on two contradictory effects of capital accumulation contemplated by Smith: a tendency of the rate of profit to fall due to the intensification of competition among capital owners; and a tendency of the rate of profit to rise due to the increase in productivity associated with the division of labour. Section 1.3 turns to David Ricardo's approach to the theory of distribution and capital

accumulation. We argue that in Ricardo the growth rate is endogenous and may fall to zero when, during capital accumulation and population growth, the rate of profit tends to fall due to diminishing returns in agriculture and the exhaustion of some natural resources. Section 1.4 deals with linear models of economic growth: the authors discussed include Robert Torrens, Karl Marx, Georg von Charasoff and John von Neumann. Section 1.5 provides a taxonomy of 'classical' cases in which the rate of profit, and thus the rate of growth, need not fall to zero. Section 1.6 discusses 'neoclassical' ideas or models of exogenous growth. Section 1.7 classifies the recent literature on the so-called 'new' growth models (NGMs) into three groups according to the route by which they try to avoid diminishing returns to capital. Section 1.8 draws some conclusions and argues that the 'new' growth theory (NGT) shares some crucial elements of the classical approach to the problem of growth and distribution.

From the point of view of method adopted in this paper it should be pointed out right at the beginning that the classical economists analysed the economic system in motion essentially in terms of a sequence of long-period positions of the economy reflecting discrete changes in the independent variables, or 'data', determining the rate of profits, rents and normal (relative) prices. These data concerned (*a*) the technical alternatives from which cost-minimizing producers can choose; (*b*) the overall level and the composition of output; and (*c*) the real wage rate of common labour (taking the available quantities of different qualities of land as given and non-depletable). There was no presumption in the classical economists that the economic system could be expected to converge to a dynamic steady state. While they illustrated certain concepts in terms of arguments that could be reinterpreted as steady-state models, they did not think that the steady state was of interest to describe and analyse real economic systems evolving in historical time. We may therefore distinguish between two concepts of 'endogenous' growth. On the one hand, we have the classical economists' view with their radical concept which sees the rate of growth at any moment and also in the long run shaped by the interaction of people between one another and with their environment. On the other hand, we have the much more narrow view entertained in the majority of contributions to new growth theory, which argue, fully in line with the Solow model, that the long run is a steady state and add, contrary to the Solow model, that the steady-state growth rate is determined from within the economic system and is not given from outside. In the following this crucial difference between the two concepts of endogeneity has to be kept in mind. In the classical economists we simply do not encounter the concept of an asymptotic growth rate of a dynamic model except as the hypothetical case of a stationary state. However, in order to have a sufficient basis for comparison between the different approaches to growth theory, old and new, we shall focus attention on models that are in fact steady-state models or ideas that are given a steady-state form.

1.2. ADAM SMITH ON GROWTH

A characteristic feature of the classical approach is the view that production involves labour, produced means of production and natural resources. In contrast to some contributions to modern growth theory none of these factors – labour, capital and land – were considered negligible other than in conceptual experiments designed 'to illustrate a principle' (Ricardo). To understand real growth processes one had to come to grips with the interrelated laws governing the growth of population, the pace of accumulation and the rate and bias of technical innovation in an environment characterized by the scarcity of natural resources. At stake was an understanding of the working of a highly complex system.

1.2.1. Capital Accumulation and the Division of Labour

Adam Smith viewed the growth process as strictly *endogenous* (see also Lowe, [1954] 1987, p. 108, and Eltis, 1984, p. 69), placing special emphasis on the impact of capital accumulation on labour productivity. He began his inquiry into the *Wealth of Nations* by stating that income per capita

> must in every nation be regulated by two different circumstances; first, by the skill, dexterity, and judgment with which its labour is generally applied; and, secondly, by the proportion between the number of those who are employed in useful labour, and that of those who are not so employed (WN I.3).

According to Smith there is no upper limit to labour productivity. This is why Smith maintained that an investigation of the growth of income per capita is first and foremost an inquiry into 'The causes of this improvement, in the productive powers of labour, and the order, according to which its produce is naturally distributed among the different ranks and conditions of men in the society' (WN I.5).

Smith's attention focused accordingly on the factors determining the growth of labour productivity, that is, the factors affecting 'the state of the skill, dexterity, and judgment with which labour is applied in any nation' (WN I.6). At this point the accumulation of capital enters into the picture, because of Smith's conviction that the key to the growth of labour productivity is the division of labour which in turn depends on the extent of

the market and thus upon capital accumulation. 'The greatest improvement in the productive powers of labour', we are told, 'seem to have been the effects of the division of labour' (WN I.i.1), both *within* given firms and industries and, even more significantly, *between* them. In his analysis in the first three chapters of book I of *The Wealth of Nations* Smith established the idea that there are *increasing returns* which are largely *external* to firms, that is, broadly compatible with the classical hypothesis of a uniform rate of profit. In the first chapter he made clear how powerful a device the division of labour is in increasing labour productivity, and analysed in some detail its major features: (*i*) the improvement of the dexterity of workers; (*ii*) the saving of time which is otherwise lost in passing from one sort of work to another; and, most importantly, (*iii*) the invention of specific machinery (see WN I.i.6–8). In the second chapter he argued that there is a certain propensity in human nature 'to truck, barter and exchange one thing for another', which appears to be rooted in 'the faculties of reason and speech', that gives occasion to the division of labour (WN I.ii.1–2). In the third chapter the argument is completed by stressing that the division of labour is limited by the extent of the market (see WN I.iii.1): a larger market generates a larger division of labour among people and, therefore, among firms, and a larger division of labour generates a larger productivity of labour for all firms.

Despite the presence of increasing returns, Smith retained the concept of a *general* rate of profit. His argument appears to be implicitly based on the hypothesis that each single firm operates at constant returns, while total production is subject to increasing returns. Even though some examples provided by Smith relate more to the division of labour within firms than to the division of labour among firms, Smith appears to be correct in sustaining that some of the activities which were originally part of the division of labour within the firm may eventually become a different 'trade' or 'business', so that the division of labour *within* the firm is but a step towards the division of labour *amongst* firms. In the example of pin making at the beginning of chapter I, Smith pointed out that 'in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades' (WN I.i.3).

Smith's analysis foreshadows the concepts of *induced* and *embodied* technical progress, *learning by doing*, and *learning by using*. The invention of new machines and the improvement of known ones is said to be originally due to the workers in the production process and 'those who had occasion to use the machines' (WN I.i.9). At a more advanced stage of society making machines 'became the business of a peculiar trade', engaging 'philosophers or men of speculation, whose trade it is, not to do any thing, but to observe every thing; and who, upon that account, are often capable of combining together the powers of the most distant and dissimilar objects'. Research and development of new industrial designs becomes 'the principal or sole trade and occupation of a particular class of citizens' (Ibidem). New technical knowledge is systematically created and economically used, with the sciences becoming more and more involved in that process. The accumulation of capital propels this process forward, opens up new markets and enlarges existing ones, increases effectual demand and is thus the main force behind economic and social development (WN V.i.e.26). Here we have a *dynamic* notion of competition, conceived of as rivalry, which anticipates in important respects the views on competition of authors such as Karl Marx and Joseph Alois Schumpeter. Smith also anticipates the following two ideas that are prominent within the 'new' growth theory literature: (1) 'new improvements of art' are generated within the economic system by specialized activities; (2) new technical knowledge is or eventually will become a *public good*, that is, non-rival and non-excludable.

Did Smith expect the endogenous growth factors to lose momentum as capital accumulates? He considered three potential limits to growth: an insufficient supply of workers, the scantiness of nature, and an erosion of the motives of accumulation. Smith saw that the scarcity and potential depletion of renewable and the depletion of exhaustible resources may constrain human productive activity and the growth of the economy (WN I.xi.i.3; see also I.xi.d). At the time when he wrote, the limits to growth deriving from nature were apparently still considered negligible.

Smith also saw no danger that the process of accumulation might come to an end because of an insufficient supply of labour and the ensuing diminishing returns to capital. He rather advocated a view which was to become prominent amongst the classical economists: the supply of labour is generated within the socio-economic system, that is, *endogenously*. He drew an analogy between the multiplication of animals and that of the inferior ranks of people (see WN I.viii.39, 40). Smith envisaged the growth of the labour force as endogenous, the determinant being the rate of capital accumulation. Real wages are higher, the more rapidly capital accumulates. As to the impact of high and rising real wages on the rate of profit, it appears that we cannot say anything definite, given Smith's opinion that 'the same cause ... which raises the wages of labour, the increase of stock, tends to increase its productive powers, and to make a smaller quantity of labour produce a greater quantity of work' (WN I.viii.57).

Surprisingly, Smith came up with a definitive answer in chapter IX of book I. His explanation of a falling tendency of the rate of profit in terms of 'competition' (WN I.ix.2) does not stand up to close examination.[1] First, since Smith commonly presupposed free competition, a fall in profitability

cannot be traced back to an intensification of competition. Second, Smith erroneously tried to carry an argument that is valid in a partial framework over to a general framework. This problem was tackled by David Ricardo.

Adam Smith explained economic growth thoroughly as an *endogenous* phenomenon. The growth rate depends on the decisions and actions of agents, especially their savings and investment behaviour, and the creativity and innovativeness they come up with in given social and historical conditions and institutional settings. Special emphasis is placed on the endogenous creation of new knowledge that can be used economically. New technical knowledge is treated as a good, which is or in the long run tends to become a public good. There are no clear and obvious limits to growth. The additional work force required in the process of accumulation is generated by that process itself: labour power is a commodity, the quantity of which is regulated by the effectual demand for it. Diminishing returns due to scarce natural resources are set aside or taken to be compensated by the increase in productivity due to the division of labour.

In the following much of the analysis will focus on whether savings (and investment) behaviour has an impact on the long-run rate of economic growth, or, in the framework of steady-state models, on the steady-state rate of growth. Accordingly, growth will be dubbed 'exogenous' if the long-run (or steady-state) rate of growth is independent of such behaviour. This narrows somewhat the scope of our investigation, because the set of purposeful decisions and actions of agents includes, but is not co-extensive with, savings decisions (see, for example, Jones, 1995; and Eicher and Turnovsky, 1999a).

1.3. DAVID RICARDO ON DIMINISHING RETURNS

Ricardo set aside what may be called *statically and dynamically increasing returns*. The beneficial effects of capital accumulation on productivity mediated through the extension of the division of labour play hardly any role in his analysis. In modern parlance, the problems of externalities which figured prominently in Smith's analysis are given only scant attention. This does not mean that Ricardo was of the opinion that they are of negligible interest. One has to recall that Ricardo explicitly subscribed to much of Smith's analysis and set himself the moderate task of correcting views of the Scotsman that he deemed wrong. These concerned especially Smith's view of the long-term trend of profitability as capital accumulates. Ricardo was keen to show that, given the real wage rate, the rate of profits cannot fall as a consequence of the 'competition of capital', as Smith had argued, but only because of diminishing returns due the scarcity of land(s). Much of Ricardo's argument therefore was developed in terms of the implicit assumption that the set of (constant returns to scale) methods of production from which cost-minimizing producers can choose, is given and constant. In such a framework the question then is how scarce natural resources affect profitability as capital accumulates. The resulting vision is reflected in what Ricardo called the ‘natural course’ of events.

As capital accumulates and population grows, and assuming the real wage rate of workers is given and constant, the rate of profit is bound to fall; due to extensive and intensive diminishing returns on land, ‘with every increased portion of capital employed on it, there will be a decreased rate of production’ (Ricardo, *Works* I, p. 98). Since profits are a residual income based on the surplus product left after the used up means of production and the wage goods in the support of workers have been deducted from the social product (net of rents), the ‘decreased rate of production’ involves a decrease in profitability. On the assumption that there are only negligible savings out of wages and rents, a falling rate of profit involves a falling rate of capital accumulation. Hence, Ricardo’s ‘natural course’ of events will necessarily end up in a stationary state. This path should not be identified with the *actual* path the economy is taking because technical progress will repeatedly offset the impact of the ‘niggardliness of nature’ on the rate of profit (see Ricardo, *Works* I, p. 120).

The assumption of a given real wage rate represents a first *logical step* in an approach to the problem of capital accumulation and income distribution which proceeds in terms of distinct analytical stages (see Garegnani, 1990). The attention focuses first on abstract and general principles which are then gradually attuned to the concrete case or specific historical circumstances under consideration. Economic theory is combined with historical analysis. Here we focus only on the first stage and set aside its historical part. The reader therefore will not be misled into thinking that in our view classical political economy is co-extensive with or can be reduced to this first stage. It reaches far beyond it.

Like Smith, Ricardo thought that saving and investment, that is, accumulation, would largely come from profits, whereas wages and rents played a negligible role. Hence, as regards the dynamism of the economy attention should focus on profitability. Assuming that the marginal propensity to accumulate out of profits, s, is given and constant, a ‘classical’ accumulation function can be formulated

$$g = \begin{cases} s(r - r_{min}) & \text{if } r \geq r_{min} \\ 0 & \text{if } r \leq r_{min} \end{cases}$$

where $r_{min} \geq 0$ is the minimum level of profitability, which, if reached, will arrest accumulation (see Ricardo, *Works* I, p. 120).

Ricardo saw the rate of accumulation as endogenous. The demand for labour is governed by the pace at which capital accumulates, the long-term supply of labour by the ‘Malthusian Law of Population’. Real wages may rise, that is, the ‘market price of labour’ may rise above the ‘natural’ wage rate. This is the case when capital accumulates rapidly, leading to an excess demand for labour. As Ricardo put it, ‘notwithstanding the tendency of wages to conform to their natural rate, their market rate may, in an improving society, for an indefinite period, be constantly above it’ (Ibidem, pp. 94–5). If such a constellation prevails for some time a ratchet effect may make itself felt: it is possible, Ricardo observed, that ‘custom renders absolute necessaries’ what in the past had been comforts or luxuries. Hence, the natural wage is driven upward by persistently high levels of the actual wage rate. Accordingly, the concept of ‘natural wage’ in Ricardo is a flexible one and must not be mistaken for a physiological minimum of subsistence.

Setting aside the complex wage dynamics in Ricardo’s theory, that is, assuming boldly a given and constant real wage rate and setting boldly the minimum rate of profit equal to zero, we may illustrate Ricardo’s view of the long-run relationship between profitability and accumulation and thus growth in a schematic way. Figure 1.1, originally used by Kaldor (1955–56), shows the marginal productivity of labour-cum-capital curve *CEGH*. It is decreasing since land is scarce: when labour-cum-capital increases, either less fertile qualities of land must be cultivated or the same qualities of land must be cultivated with processes which require less land per unit of product, but are more costly in terms of labour-cum-capital. Let the real wage rate equal *OW*. Then, if the amount of labour-cum-capital applied is L_1, the area $OCEL_1$ gives the product, $OWDL_1$ gives total capital employed, and *BCE* total rent.

Profit is determined as a residual and corresponds to the rectangle *WBED*. As a consequence, the *rate* of profit can be determined as the ratio of the areas of two rectangles which have the same base and, therefore, it equals the ratio *WB/OW*. Let us now consider the case in which the amount of labour-cum-capital is larger, that is, L_2. Then $OCGL_2$ gives the product, $OWFL_2$ the capital, *ACG* the rent, and *WAGF* the profits. The rate of profit has fallen to *WA/OW*. Obviously, if a positive profit rate implies a positive growth rate, the economy will expand until labour-cum-capital has reached the level $\bar{L}$. At that point the profit rate is equal to zero and so is the growth rate. The system has arrived at the so-called stationary state: growth has come to an end because profitability has.

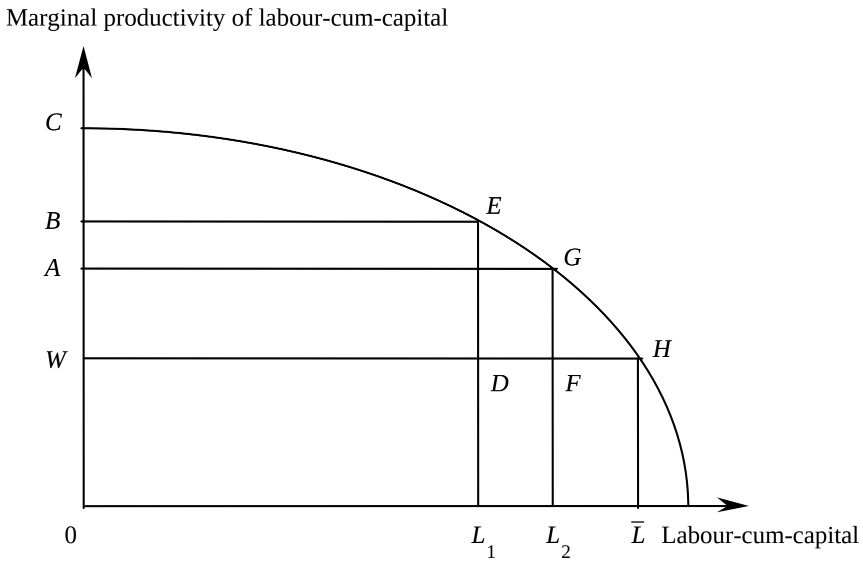


Figure 1.1: The one-commodity Ricardian model with land as an indispensable resource

For both Smith and Ricardo the required size of the work force is essentially generated by the accumulation process itself. In other words, labour power is treated as a kind of producible commodity. It differs from other commodities in that it is not produced in a capitalistic way in a special industry on a par with other industries, but is the result of the interplay between the growth of the working population and socioeconomic conditions. In the most simple and abstract conceptualization possible, labour power is seen to be in elastic supply at a given real wage basket. Increasing the number of baskets available in the support of workers involves a proportional increase of the work force. In this view the rate of growth of labour supply adjusts to any given rate of growth of labour demand without necessitating a variation in the real wage rate.

In a more sophisticated conceptualization, higher rates of growth of labour supply presuppose higher levels of the real wage rate. But the basic logic remains the same: in normal conditions the pace at which capital accumulates regulates the pace at which labour, a *non-accumulable* factor of production, grows. Thus labour cannot put a limit to growth because it is generated within the growth process. The only limit to growth can come from other non-accumulable factors of production. In other words, there is only endogenous growth in Ricardo. This growth is bound to lose momentum as the system hits its natural barriers, especially as soon as

extensive and intensive diminishing returns make themselves felt and are not counteracted by sufficient technical progress. This also shows that it is not a necessary condition, for a theory to be considered a theory of endogenous growth, that it assumes some kind of increasing returns. This becomes clear in Section 1.4 below. Interestingly, its main message was anticipated by Ricardo.

Ricardo contemplated the implications for income distribution and the rate of expansion of the economic system in the hypothetical case in which land of the best quality is available in abundance. In one place he wrote:

> Profits do not necessarily fall with the increase of the quantity of capital because the demand for capital is infinite and is governed by the same law as population itself. They are both checked by the rise in the price of food, and the consequent increase in the price of labour. If there were no such rise, what could prevent population and capital from increasing without limit? (Ricardo, *Works* VI, p. 301)

If land of the best quality were available in abundance it would be a free good and no rent would be paid for its use. In this case the curve of the graph showing the marginal productivity of labour-cum-capital would be a horizontal line and the rate of profit would be constant whatever the amount of labour-cum-capital employed. As a consequence, other things being equal, the growth rate would also be constant: the system could grow for ever at a rate that equals the given rate of profit times the propensity to accumulate. As the passage from Ricardo's *Works* just quoted shows, Ricardo was perfectly aware of this implication.

1.4. LINEAR CLASSICAL MODELS OF PRODUCTION

Central elements of classical analysis are the concept of production as a circular flow and the related concept of *surplus* product left after the wage goods and what is necessary for the replacement of the used up means of production have been deducted from the annual output. This surplus can be consumed or accumulated. With constant returns to scale and setting aside the problem of scarce natural resources, the notion of an economy expanding at a constant rate of growth was close at hand. In this section we shall mention some contributions to what may be called linear growth theory with a classical flavour.

Robert Torrens in his *Essay on the External Corn Trade* clarified that the concept of surplus provides the key to an explanation of the *rate* of profit. Growth in the model by Torrens is both linear and endogenous; the rate of growth depends on the general rate of profit and the propensity to accumulate. The same can be said of Marx's theory of expanded reproduction in chapter 21 of volume II of *Capital* (Marx, [1885] 1956).[2] There Marx studied the conditions under which the system is capable of reproducing itself on an upward spiralling level. The expansion of the economy at an endogenously determined rate of growth is possible. This rate depends on the proportion of the surplus value ploughed back into the productive system to increase the scale of operation. Marx stressed that the accumulation of capital is 'an element *immanent* in the capitalist process of production' (Ibidem, p. 497; emphasis added). For, 'the aim and compelling motive of capitalist production' is 'the snatching of surplus-value and its capitalisation, i.e., accumulation' (Ibidem, p. 507).

The Russian mathematician Georg von Charasoff elaborated on Marx's analysis and was possibly the first to provide a clear statement of the fundamental duality relationship between the system of prices and the rate of profit on the one hand, and the system of quantities and the rate of growth on the other (see Charasoff, 1910). He developed his main argument within the framework of an interdependent model of (single) production exhibiting all the properties of the later input–output model, and which is fully specified in terms of use values (rather than labour values as in the case of Marx) and labour needed per unit of output.

The most sophisticated linear model of endogenous growth was elaborated by John von Neumann (1945) in a paper first published in German in 1937 and then translated into English in 1945. In it von Neumann assumed there are *n* goods produced by *m* constant returns–to–scale production processes. There is a problem of the choice of technique which consists in establishing which processes will actually be used and which not, being 'unprofitable'. Von Neumann (1945, pp. 1–2) took the real wage rate, consisting of the 'necessities of life', to be given and paid at the beginning of the uniform period of production, that is, he considered wages as part of the capital advanced and thus as part of the *physical real costs of production*. In addition, he assumed 'that all income in excess of necessities of life will be reinvested'. In von Neumann's model the rate of growth is determined *endogenously*.[3] He set aside the problem of scarcity of all non-accumulable factors of production: while all primary factors other than labour (that is, all natural resources) were taken to be available at whichever amount was needed at zero price, labour was assumed to be available at the required amount at a given real wage rate.

1.5. A CLASSIFICATION OF CASES

We can now classify some broad cases in which the rate of profit, and therefore the rate of growth, does not fall to zero. There is perpetual growth provided that the premises underlying the different cases hold infinitely. It will be seen that while the cases discussed are all derived from a classical framework of the analysis as it was developed by Adam Smith and David Ricardo, the cases exhibit some striking similarities to the types of NGMs discussed in Section 1.7.

1.5.1. Constant Returns to Capital

As we have seen, the main ingredient to obtain a stationary state in the Ricardian model is the existence of land available in limited supply. If land were not needed as an input or if land of the best quality were available in abundance, then the graph giving the marginal productivity of labour-cum-capital would be a horizontal line and therefore the rate of profit would be constant whatever the amount of labour-cum-capital. As a consequence, the growth rate would also be constant.

1.5.2. Backstop Technology

To assume that land is not useful in production or that it is available in given quality and unlimited quantity is unnecessarily restrictive. It is enough to assume that 'land', although useful in production, is not indispensable. In other words, there is a technology that allows the production of the commodity without any 'land' input. With continuous substitution between labour-cum-capital and land, the marginal productivity of labour-cum-capital would be continuously decreasing, but it would be bounded from below. With respect to the case depicted in Figure 1.1, the CEGH curve is decreasing, as in Figure 1.1, but the concavity is reversed and there is a horizontal asymptote. In this case the profit rate and thus the growth rate are falling, but they could never fall below certain positive levels. The system would grow indefinitely at a rate of growth that asymptotically approaches the product of the given saving rate times the value of the (lower) boundary of the profit rate.

1.5.3. Increasing Returns to Capital

The final case is that of increasing returns to labour-cum-capital as was discussed, following Adam Smith, by Allyn Young (1928) and Nicholas Kaldor (1957 and 1966). Taking the wage rate as given and constant, the rate of profit and the rate of growth will rise as more labour-cum-capital is employed. To preserve the notion of a uniform rate of profit, it is necessary to assume that the increasing returns are *external* to the firm and exclusively connected with the expansion of the market as a whole and the social division of labour. This implies that while in the case of decreasing returns due to the scarcity of land the product was given by the area under the marginal productivity curve, now the product associated with any given amount of labour-cum-capital is larger than or equal to that amount multiplied by the corresponding level of output per unit of labour-cum-capital. In any case, the sum of profits and wages equals the product of the given amount of labour-cum-capital multiplied by the corresponding level of output per unit of labour-cum-capital.[4] As a consequence, the product is larger than the area under the marginal productivity curve. The cases of decreasing and increasing returns are therefore not symmetrical: with increasing returns a rising real wage rate need not involve a falling general rate of profit.

1.6. MODELS OF EXOGENOUS GROWTH

The marginalist or 'neoclassical' school of economic thought seeks to explain income distribution in a symmetrical way via the relative scarcities of the factors of production, labour, 'capital,' and land. Interestingly, the idea of *exogenous* growth which classical theory did *not* entertain is the starting point of important early works in the marginalist tradition.

1.6.1. Alfred Marshall and Gustav Cassel

The idea of an economic system growing exclusively because some exogenous factors make it grow has variously been put forward in the history of economic thought as a standard of comparison. For example, in chapter V of book V of his *Principles*, first published in 1890, Alfred Marshall ([1890] 1977, p. 305) introduced the 'famous fiction of the "Stationary state" ... to contrast the results which would be found there with those in the modern world'. By relaxing one after another of the rigid assumptions defining the stationary state, Marshall sought to get gradually closer to the 'actual conditions of life'. The first relaxation concerned the premise of a constant (working) population:

> The Stationary state has just been taken to be one in which population is stationary. But nearly all its distinctive features may be exhibited in a place where population and wealth are both growing, provided they are growing at about the same rate, and there is no scarcity of land: and provided also the methods of production and the conditions of trade change but little; and above all, where the character of man himself is a constant quantity. For in such a state by far the most important conditions of production and consumption, of exchange and distribution will remain of the same quality, and in the same general relations to one another, though they are all increasing in volume (Ibidem, p. 306).

The resulting economic system grows at a constant rate which equals the exogenous rate of growth of population.[5] Income distribution and relative prices are the same as in the stationary economy. In modern parlance: the system expands along a steady-state growth path.

We encounter essentially the same idea in Gustav Cassel's ([1918] 1932) *Theory of Social Economy*. The model of exogenous growth delineated by Cassel can be considered the proximate starting point of the development of neoclassical growth theory. In chapter IV of book I of the treatise Cassel presented two models, one of a stationary economy, the other of an economy growing along a steady-state path.

In his first model Cassel assumed that there are *z* (primary) factors of production. The quantities of these resources and thus the amounts of services provided by them are taken to be in given supply. General equilibrium is characterized by equality of supply and demand for each factor service and for each good produced, and equality of the price of a good and its cost of production. The resulting sets of equations constitute what is known as the 'Walras–Cassel model' (Dorfman, Samuelson and Solow, 1958, p. 346). It satisfies the then applied criterion of completeness: there are as many equations as there are unknowns to be ascertained.

Cassel (1932, pp. 152–3) then turned to the model of a uniformly progressing economy. Although described only verbally, he introduced the model in the following way:

> We must now take into consideration the society which is progressing at a uniform rate. In it, the quantities of the factors of production which are available in each period ... are subject to a uniform increase. We shall represent by [*g*] the fixed rate of this increase, and of the uniform progress of the society generally.

In Cassel's view this generalization to the case of an economy growing at an exogenously given and constant rate does not cause substantial problems. The previously developed set of equations can easily be adapted appropriately, 'so that the whole pricing problem is solved'. Cassel thus arrived at basically the same result as Marshall.

1.6.2. Robert Solow, Trevor Swan and James Meade

The neoclassical growth models of the 1950s and early 1960s differ from the growth version of the Walras–Cassel model in five important respects:

1. they are macro-models with only one produced good which could be used both as a consumption good and as a capital good;
2. the number of primary factors of production is reduced to one, homogeneous labour (as in Solow, 1956 and 1963; Swan, 1956), or two, homogeneous labour and homogeneous land (as in Swan, 1956; Meade, 1961);
3. the all-purpose good is produced by means of labour, capital, that is, the good itself, and possibly land;
4. there is a choice of technique, where technical alternatives are given by a macroeconomic production function, which is homogeneous of degree one with positive and decreasing marginal productivities with respect to each factor of production; and
5. planned saving, which is taken to be equal to planned investment at all times, is proportional to net income, that is, a ‘Keynesian’ saving function is assumed.

Focusing attention on the models with a single primary factor (labour), in steady-state equilibrium

$$sf(k) = gk,$$

where s is the (marginal and average) propensity to save, $f(k)$ is the production function per unit of labour or *per capita*, k is the capital–labour ratio (where labour is measured in terms of efficiency units), and g is the steady-state growth rate of capital (and labour, and income etc.). In steady-state equilibrium output expands exactly as the exogenous factors make it grow. Note that assuming $s > 0$ presupposes that the exogenous factors are growing at some positive rate. In these models the steady-state rate of growth is exogenous. Outside steady-state equilibrium the rate of growth can be shown to depend also on the behavioural parameter of the system, that is, the propensity to save (and invest), but that parameter plays no role in determining the long-term rate of growth.

While these models are aptly described as models of *exogenous* growth, they can also be described as models of *endogenous* profitability. Since in the one-good framework adopted by the authors under consideration the rate of profit r equals the marginal productivity of capital,

$$r = f'(k),$$

the two equations are able to determine a relationship between the rate of profit and the steady-state rate of growth. The following section shows that the NGMs essentially reverse what is endogenous and what is exogenous. In other words, without over-exaggeration they can be called models of *endogenous* growth and *exogenous* profitability.

1.7. THE ‘NEW’ MODELS OF ENDOGENOUS GROWTH

One of the key properties of the NGMs emphasized by their advocates is the limitation of diminishing returns to capital. The first generation of NGMs defined the confines within which subsequent contributions to NGT were carried out. The attention focuses on the mechanism that prevents the returns to capital from falling (below a certain level).[6]

1.7.1. Constant Returns to Capital

The first class of models set aside all non-accumulable factors of production such as labour and land and assume that all inputs in production are accumulable, that is, ‘capital’ of some kind. The simplest version of this class is the so-called ‘AK model’, which assumes that there is a linear relationship between total output, Y, and a single factor capital, K, both consisting of the *same* commodity:

$$Y = AK, \tag{1}$$

where $1/A$ is the amount of that commodity required to produce one unit of itself. Because of the linear form of the aggregate production function, these models are also known as ‘linear models’. This model is immediately recognized as the model dealt with in Subsection 1.5.1. The rate of return on capital r is given by

$$r + \delta = \frac{Y}{K} = A, \tag{2}$$

where δ is the exogenously given rate of depreciation. There is a large variety of models of this type in the literature. In the two-sector version in Rebelo (1991) it is assumed that the capital good sector produces the capital good by its own means and nothing else. It is also assumed that there is only one method of production to produce the capital good. Therefore, *the rate of profit is determined by technology alone*. Then the saving-investment mechanism jointly with the assumption of a uniform rate of growth, that is, a steady-state equilibrium, determines a relationship between the growth rate, g, and the rate of profit, r. Rebelo (1991, pp. 504 and 506) obtains either

$$g = \frac{A - \delta - \rho}{\sigma} = \frac{r - \rho}{\sigma}, \tag{3}$$

or

$$g = (A - \delta)s = sr. \tag{4}$$

Equation (3) is obtained when savings are determined on the assumption that there is an immortal representative agent maximizing the following inter-temporal utility function

$$\int_0^\infty e^{-\rho t}\frac{1}{1-\sigma}\left[c(t)^{1-\sigma}-1\right]\mathrm{d}t\,,$$

subject to constraint (1), where ρ is the discount rate, or rate of time preference, and $1/\sigma$ is the elasticity of substitution between present and future consumption ($1 \neq \sigma > 0$), and where $Y = c(t) + \dot{K}$. Equation (4) is obtained when the average propensity to save s is given. Hence, in this model the rate of profit is determined by technology alone and the saving-investment mechanism determines the growth rate.

King and Rebelo (1990) essentially followed the same avenue. Instead of one kind of 'capital' they assumed that there are two kinds, real capital and human capital, both of which are accumulable. There are two lines of production, one for the social product and the real capital, which consist of quantities of the same commodity, and one for human capital. The production functions relating to the two kinds of capital are assumed to be homogeneous of degree one and strictly concave. There are no diminishing returns to (composite) capital for the reason that there is no non-accumulable factor such as simple or unskilled labour that enters into the production of the accumulable factors, investment goods and human capital. As in Rebelo's model the rate of profit is uniquely determined by the technology (and the maximization of profits which, because of the Non-substitution Theorem,[7] implies that only one technique can be used in the long run); the growth rate of the system is then endogenously determined by the saving–investment equation. The greater the propensities to accumulate human and physical capital, the higher is the growth rate.

1.7.2. Returns to Capital Bounded from Below

The second class of models preserve the dualism of accumulable and non-accumulable factors but restrict the impact of an accumulation of the former on their returns by modification of the aggregate production function. Jones and Manuelli (1990), for example, allow for both labour and capital and even assume convex technology, as the Solow model does. However, convex technology requires only that the marginal product of capital is a decreasing function of its stock, not that it vanishes as the amount of capital per worker tends towards infinity. Jones and Manuelli assume that

$$h(k) \geq bk, \quad \text{each } k \geq 0,$$

where $h(k)$ is the per capita production function and b is a positive constant. The special case contemplated by them is

$$h(k) = f(k) + bk, \tag{5}$$

where $f(k)$ is the conventional per capita production function. As capital accumulates and the capital–labour ratio rises, the marginal product of capital will fall, approaching asymptotically b, its lower boundary. With a given propensity to save, s, and assuming capital never wears out, the steady-state growth rate g is endogenously determined: $g = sb$. Assuming, on the contrary, intertemporal utility maximization, the rate of growth is positive provided the technical parameter b is larger than the rate of time preference ρ. In the case in which it is larger, the steady-state rate of growth is given by equation (3) with $r = b$.

It may be easily recognized that the difference between the model of Jones and Manuelli (1990) and that of Rebelo (1991) is the same as that existing between the case dealt with in Subsection 1.5.1 and that dealt with in Subsection 1.5.2.

1.7.3. Factors Counteracting Diminishing Returns to Capital

Finally, there is a large class of models contemplating various factors counteracting any diminishing tendency of returns to capital. Here we shall be concerned only with the following two sub-classes: human capital formation and knowledge accumulation. In both kinds of models *positive external effects* play an important part; they offset any fall in the marginal product of capital.

A. *Human capital formation*

Models of the first sub-class attempt to formalize the role of human capital formation in the process of growth. Elaborating on some ideas of Uzawa (1965), Lucas (1988) assumed that agents have a choice between two ways of spending their (non-leisure) time: to contribute to current production or to accumulate human capital. With the accumulation of human capital there is said to be associated an externality: the more human capital society as a whole has accumulated, the more productive each single member will be. This is reflected in the following macroeconomic production function

$$Y = AK^{\beta}(uhN)^{1-\beta}h^{*\gamma}, \tag{6}$$

where the labour input consists of the number of workers, N, times the fraction of time spent working, u, times h which gives the labour input in efficiency units. Finally, there is the term h^*. This is designed to represent the externality. The single agent takes h^* as given in his or her optimizing.

However, for society as a whole the accumulation of human capital increases output both directly and indirectly, that is, through the externality.

Lucas's conceptualization of the process by means of which human capital is built up is the following:

$$\dot{h} = \upsilon h(1 - u), \tag{7}$$

where υ is a positive constant. (Note that equation (7) can be interpreted as a 'production function' of human capital.)

Interestingly, it can be shown that if the above-mentioned externality is *not* present, that is, if γ in equation (6) equals zero, and therefore returns to scale are constant and, as a consequence, the Non-substitution Theorem holds, endogenous growth in Lucas's model is obtained in essentially the same way as in the models by Rebelo (1991) and King and Rebelo (1990): the rate of profit is determined by technology and profit maximization alone; and for the predetermined level of the rate of profit the saving–investment mechanism determines the rate of growth. Yet, as Lucas himself pointed out, endogenous growth is positive *independent* of the fact that there is the above-mentioned externality, that is, independent of the fact that γ is positive.[8] *Therefore, while complicating the picture, increasing returns do not add substantially to it: growth is endogenous even if returns to scale are constant.* If returns to scale are not constant then the Non-substitution Theorem does not apply, implying that neither the competitive technique nor the associated rate of profit are determined by technical alternatives and profit maximization alone. Nevertheless, these two factors still determine, in steady states, a relationship between the rate of profit and the rate of growth. This relationship together with the relationship between the same rates obtained from the saving–investment mechanism determines both variables. Although the analysis is more complex, essentially the same mechanism applies as in the models dealt with in Subsection 1.7.1.

B. *Technical change*

Models of the second sub-class attempt to portray technological change as generated endogenously. The proximate starting point for this kind of model was Arrow's (1962) paper on 'learning by doing'. Romer (1986) focuses on the role of a single state variable called 'knowledge' or 'information' and assumes that the information contained in inventions and discoveries has the property of being available to anybody to make use of it at the same time. In other words, information is considered essentially a non-rival good. Yet, it need not be totally non-excludable, that is, it can be monopolized at least for some time. It is around the two different aspects of publicness – non-rivalry and non-excludability – that the argument revolves. Discoveries are made in research and development departments of firms. This requires that resources

be withheld from producing current output. The basic idea of Romer's (1986, p. 1015) model is 'that there is a trade-off between consumption today and knowledge that can be used to produce more consumption tomorrow'. He formalizes this idea in terms of a 'research technology' that produces 'knowledge' from forgone consumption. Knowledge is assumed to be cardinally measurable and not to depreciate: it is like perennial capital.

Romer stipulates a research technology that is concave and homogeneous of degree one,

$$\dot{k}_i = G(I_i, k_i), \tag{8}$$

where I_i is an amount of forgone consumption in research by firm i and k_i is the firm's current stock of knowledge. (Note that the forgone consumption good is a capital good utilized in the production of 'knowledge'.) The production function of the consumption good relative to firm i is

$$Y_i = F(k_i, K, \mathbf{x}_i), \tag{9}$$

where K is the accumulated stock of knowledge in the economy as a whole and $\mathbf{x}_i$ are all inputs different from knowledge. The function is taken to be homogeneous of degree one in k_i and $\mathbf{x}_i$ and homogeneous of a degree greater than one in k_i and K. Romer (1986, p. 1019) assumes that 'factors other than knowledge are in fixed supply'. This implies that 'knowledge' is the only *capital good* utilized in the production of the consumption good. Spillovers from private research and development activities increase the public stock of knowledge K.

Assuming, contrary to Romer, that the above production function (9) is homogeneous of degree one in k_i and K involves a constant marginal product of capital: the diminishing returns to k_i are exactly offset by the external improvements in technology associated with capital accumulation. In this case it can be shown that, similar to the NGMs previously dealt with, the rate of profit is determined by technology and profit maximization alone, provided, as is assumed by Romer, that the ratio K/k_i equals the (given) number of firms. The saving–investment relation then determines endogenously the growth rate. Once again, endogenous growth does not depend on an assumption about increasing returns with regard to accumulable factors. Growth would be no more endogenous if increasing returns were to be assumed: such an assumption would only render the analysis a good deal more complicated. In particular, a steady-state equilibrium does not exist, and in order for an equilibrium to exist the marginal product of capital must be bounded from above. This is effected by Romer in terms of an *ad hoc* assumption regarding equation (8) (Ibidem, p. 1019). This assumption is not different from the one used in drawing

Figure 4, where the marginal product of corn is shown to be increasing with the scale of production, but is bounded from above.

1.8. CONCLUSION

The NGMs revolve around a few simple and rather obvious ideas which have been anticipated by earlier economists, most notably Adam Smith and David Ricardo. Many of the interesting aspects of the NGMs are related to the classical perspective that their authors (often unwittingly) take on the problem of growth, whereas some of their shortcomings derive from the lack of solutions to the problems of the neoclassical theory of growth which were put into sharp relief during the 1960s and 1970s. It has also been hinted that in some non-neoclassical approaches to the theory of accumulation and growth, the endogeneity of the growth rate has always been taken for granted. A brief look into the history of economic thought shows that from Adam Smith via David Ricardo, Robert Torrens, Thomas Robert Malthus, Karl Marx up to John von Neumann both the equilibrium and the actual rate of capital accumulation and thus both the equilibrium and the actual rate of growth of output as a whole were seen to depend on agents' behaviour, that is, endogenously determined. In this regard there is indeed nothing new under the sun.

NOTES

* We should like to thank Mauro Caminati for valuable comments and suggestions. All remaining errors and omissions are entirely our responsibility.

1. For an interesting different view placing special emphasis on Malthus's interpretation of Smith according to which Smith had ruled out constant and diminishing returns, see Negishi (1993).
2. In Marx's analysis, this theory is only a logical step toward a proper theory of accumulation. Here we cannot deal with the latter and Marx's 'law' of a falling tendency of the rate of profits in Volume III of *Capital*. Here Marx argues that a tendency of the real wage rate toward a socially and historically defined subsistence level is not due to a population mechanism, but due to the presence of an 'industrial reserve army of the unemployed', which is continually filled and re-filled by labour–saving technical progress.
3. This is one of the reasons why the conventional interpretation of that model as belonging to the tradition established by the so-called 'Walras–Cassel model' cannot be sustained (see Kurz and Salvadori, 1993). Cassel (1932) took as exogenously given the rates of growth of all primary factors and assumed their continuous full employment (see Section 1.6 below). Von Neumann never made this assumption.
4. Let $x = f(L, L^*)$ be the product of the last unit of labour-cum-capital, when L represents the amount of labour-cum-capital employed and the division of labour is artificially kept fixed at the level appropriate when the amount of labour-cum-capital employed is L^*. Obviously, $f(L, L^*)$ as a function of L alone is either decreasing (if land is scarce) or constant (if land is not scarce). The product at L^* equals $\int_0^{L^*} f(L, L^*)\mathrm{d}L$, i.e., the area under the curve $f(L, L^*)$ in the range $[0, L^*]$. If $(\partial f/\partial L^*) > -(\partial f/\partial L)$ for $L^* = L$, then the curve $x = f(L, L^*)$ is increasing, but the product is, as stated in the text, larger than or equal to the sum of profits and wages, which equals the product of the given amount of labour-cum-capital multiplied by the corresponding level of output per unit of labour-cum-capital.

5. It should be noted that Marshall (1977, book IV, ch. IV) saw reason to suppose that the growth of population depended, among other things, on socioeconomic factors and thus could not sensibly be treated, other than in an initial step of the analysis, as exogenous.
6. For a more detailed treatment of these models, see Kurz and Salvadori (1998b).
7. We need a special case of the Non-substitution Theorem, because no primary factor (or a primary factor with a zero remuneration) is assumed; see Kurz and Salvadori (1994).
8. For a demonstration of this, see Kurz and Salvadori (1998b).

2. The structure of growth models: a comparative survey*

Antonio D'Agata and Giuseppe Freni

2.1. INTRODUCTION

The basic problem of growth theory is to describe the behaviour of an expanding economy over time. The more traditional way to conceive growth is to consider it as due to the accumulation of capital. In its attempt to construct a theory of growth (see e.g. Solow, 1956), neoclassical economics sought to extend its static theory of distribution to a dynamic context, and in order to succeed in this attempt it had to assume decreasing returns with respect to the accumulated factor (see e.g. Bertola, 1993, 1994). This assumption results in the accumulation process having only transitory effects on the rate of growth, whose long-run behaviour therefore remains unexplained within the model and is characterised by the constancy of the capital/labour and product/labour ratios. As a consequence, empirically relevant examples of permanent growth, like sustained increase in the per capita stock of capital, are attributed to the 'compensating influence of residual factors that have been assumed away in the model' (Kaldor, 1961, p. 177).

The properties of the neoclassical growth theory have always been questioned not only on empirical but also on theoretical grounds. One of the main criticisms has been that the rate of growth of economies should depend upon the thriftiness of the economy and that technical change should be the outcome of intentional decisions of economic agents. The recent literature on the endogenous growth theory has been successful in dealing with such criticisms and has been able to construct a variety of models in which the rate of growth depends upon the saving decision of households and/or technological change is the intentional or unintentional outcome of the maximising behaviour of agents.

Kurz and Salvadori (1998b; 1999) and other authors have pointed out that the endogenous growth theory represents a substantial break with respect to the neoclassical theory by recovering various aspects of the classical view of

the economic process. In fact, while the neoclassical theory of factor–income distribution is hardly consistent with endogenous growth (Frankel, 1962), classical economists, basing their distribution theory on non-economic elements and assuming full reproducibility of means of production and constant returns, constructed a theory of growth which was able to ensure the possibility of persistent (endogenous) growth.

While the work of Kurz and Salvadori is helpful in highlighting the continuity of the endogenous growth theory with the classical tradition, its scope is too restricted to allow full appreciation of the features of each theory in terms of common and idiosyncratic elements. On the other hand, a complete idea of the basic structure of each theory is important for possible cross-fertilisation between different streams of analysis. The aim of this paper is to provide a first attempt in this direction by examining in a very simple way the structure of the most important growth models. While we accept Kurz and Salvadori's idea that the endogenous growth theory can be considered as retrieving the classical view as far as distribution and the sources of growth are concerned, we shall try to point out the existence of common aspects and specific features of the models considered. To be more specific we shall emphasise that there is a continuity between classical growth theory, neoclassical growth theory and endogenous growth theory as far as the relationship between savings and investment is concerned, in that all these models conceive the investment decision as always co-ordinated with the saving decision while, as is well known, the Keynesian models propose a radically different view in which investment determines savings and thus the theory of growth proves intrinsically connected with the theory of the business cycle. However, as we will point out, there are elements of continuity between the classical tradition, the Keynesian one and the endogenous growth theory with regard to the view concerning the adaptation of the rate of growth of population and that of investment.

The paper is organised as follows. In the next section the early contributions to growth theory by classical economists and von Neumann are illustrated. In Section 2.3 the Keynesian tradition will be considered. Section 2.4 deals with various streams of neoclassical growth theory, from the Solow model to the discounted Ramsey models of Cass (1965) and Koopmans (1965). In Section 2.5 we shall review the Lucas and the 1990 Romer models of endogenous growth. In Section 2.6 we shall summarize the main aspects of each model. We seek to emphasize that our analysis is not exhaustive, and several models are not considered. Among others, all overlapping generation models and models with altruism are excluded from analysis.

2.2. ACCUMULATION-ORIENTED MODELS

2.2.1. Classical Models of Economic Growth

Classical economists centred their attention on the economic growth of nations and explained this phenomenon through a theory based upon the class structure of the capitalist economy. They identify three classes, workers, capitalists and landowners, which have their own specific role in the economic process. Workers own labour and sell it on the labour market for a subsistence wage. Landowners rent their land to capitalists in order to obtain rent. Capitalists own the produced means of production and organise production by employing labour and renting land. Profits are their income. The behaviour of all these agents is governed by their attempt to get as much as possible from the resources they own.

As far as the use of the income is concerned, the usual interpretation of classical economists is that they conceive each class as being characterised by specific behaviour (Kaldor, 1961, p. 180):[1] while workers and landowners substantially consume all the income they get, capitalists save and invest essentially the entire amount of profit. However, while workers buy mainly subsistence goods, due to the low level of wages, landowners buy mainly luxury goods. An important aspect of the classical view on the consumption decision is that allocation of income is not determined by preferences or by the type of income earned but mainly by the social group to which those who receive the income belong. The sociological analysis of the rise of capitalism carried out by Smith in Book III of *The Wealth of Nations,* makes this position very clear by pointing out that capitalism arises from the emergence of a class, the 'merchants and artificers', which 'acted merely from a view to their own interest', by contrast '[T]o gratify the most childish vanity was the sole motive of the great proprietors' (WN III.iv.17).[2] Capitalists are the class which use their riches in order to improve their condition by accumulating them. The mechanism by which savings are transformed into investment can be direct, if the savers are also entrepreneurs, or indirect, through the capital market, if savers are not entrepreneurs. In any case classical economists accept the view that all savings are transformed into investment (see e.g. Hagemann, 1998). However, while classical economists accept the idea that savings can be equalised to investments through the capital market, it can be hardly said that they share the neoclassical view according to which this equalisation is due to the adjustment of the interest rate (Hahn and Matthews, 1964, pp. 12–15). In fact, for example, for Smith and Ricardo the interest rate is determined by the rate of profits (WN I.vi.18, *Works*, I, pp. 363–4), while the equalisation between savings and investment is ensured by the fact that 'the demand for capital is infinite' at the current rate of interest (*Works*,

VI, p. 301). Moreover, the rate of profit is determined by the wage rate, which in turn depends upon the conditions of the labour market. Classical economists held that the rate of growth of population depends upon the wage rate: a wage rate higher than the natural one, i.e. that which maintains constant population (*Works*, I, p. 91), yields a population increase (WN I.viii.34–42, *Works*, Ibidem). More specifically the higher the wage rate, the higher the rate of growth of population (WN I.viii.22).[3]

Hence Smith, anticipating Malthus, conceives the demographic law as a technological rule for producing labour as any good: '… the demand for men, like that for any other commodity, necessarily regulates the production of men; quickens it when it goes on too slowly, and stops it when it advances too fast' (WN I.viii.40).

Within the classical framework it is possible to consider two kinds of models. The first is the Ricardian model, which emphasises the tendency towards the stationary state due to the existence of scarce natural resources, the second model is rooted in the Smithian and Marxian tradition which emphasises the progressive nature of economic growth. Both models are commonly interpreted as being characterised by the assumptions that capital and labour are employed in fixed proportions, capitalists anticipate wages that are entirely consumed by workers, landowners consume all rents and capitalists/entrepreneurs invest all profits. While the existing amount(s) of land(s) is (are) constant, labour supply is constant only in the short run, whereas in the long run it is infinitely elastic at the natural wage rate w^*. This means that the supply of labour can be eventually increased (or decreased) indefinitely at the wage rate w^*, although the Malthusian law mentioned above regulates the rate of increase of population during any transitional phase.

To illustrate the structure of the basic classical model, we consider a growing economy at generic time t. At the beginning of this period, the economy is endowed with a set of available methods of production, M_t, with a given amount of physical capital, X_t, with a given amount of labour, N_t, and with a given amount of natural resources (land), T (T is a vector if land is of different quality). The amount of land is assumed to be constant, while the other magnitudes can change over time. Set M_t must be interpreted as the set of methods describing the output per worker that can be obtained on the different pieces of land or on the same piece of land and that are available in the economy. Every agent wants to employ all the resources he/she owns; in particular, capitalists end up employing the whole capital X_t. Competition among capitalists (workers) equalizes the rates of profits (the wage rate) on different industries, while competition among landowners drives the rent on idle lands to zero. For the sake of simplicity, following Smith,we assume that the economy is growing and that the wage rate is at such a level as to ensure a growth of population equal to the growth of demand for labour.[4]

For a given amount of capital X_t, therefore, the wage rate w_t determines the set of cost-minimising methods of production, M_t^*, and the rate of profits. By activating the cost-minimising methods of production, capitalists determine the allocation of total capital between physical capital (K_t) and wage goods anticipated to workers (W_t), and the decisions concerning the demand for labour (L_t). The net production $Y_t - X_t$ is determined as well and is devoted to pay rents and profits. Rents are entirely spent in consumption while profits are entirely saved and reinvested. At the beginning of time t+1, the investment I_t determines the stock of capital X_{t+1}, the available technology, M_{t+1}, may be different from the technology available the preceding period,[5] the supply of labour is changed according to an exogenously given Malthusian rule, while the amount of natural resources remains unchanged, by assumption.

The Ricardian model can be considered a particular version of the basic classical model, and it is characterised by the further assumptions that the set of available methods is constant over time and that land scarcity that goes with capital accumulation pushes output per worker down below what it is needed for reproduction. By contrast, the Smithian–Marxian model is characterised by non-decreasing productivity of labour and capital.6 This is justified by the fact that Smith and Marx do not lay any emphasis on the role of natural resources as a factor limiting growth. The Ricardian and the Smithian-Marxian model have been analysed by Kurz and Salvadori (this volume Ch.1) to which we refer.

Here, we focus upon the short run adjustment towards the steady state path with full employment in the Smithian-Marxian theory. The logic of this process can be intuitively grasped by means of Figure 2.1.[7] Let us assume that there is only one good produced by itself as circulating capital and by labour under constant returns, and that a single method of production is available. Without loss of generality we can now suppose that the given method uses v units of capital and 1 unit of labour per unit of output. Let Y_t be the production level of period t. Then $Y_t = \min(K_t/v, L_t)$, where K_t is the amount of circulating capital employed at time t and L_t is labour employed at time t ($W_t = w_t L_t$, where w_t is the wage rate at time t, is therefore the wages capital anticipated to workers at time t). It is assumed that the fraction s of profits is saved. Time is discrete. With X_t representing the total capital employed at time t, if the wage rate is positive, then from the equilibrium conditions $X_t = K_t + w_t L_t$, $Y_t = K_t/v = L_t$ one obtains $Y_t(w_t) = L_t(w_t) = X_t/(v+w_t)$, $K_t(w_t) = vX_t/(v+w_t)$, $W_t(w_t) = w_t X_t/(v+w_t)$. The first relation associates to each stock of total capital the demand of labour (or, equivalently, the level of production), the second and the third associate to each level of total capital

the circulating capital and the wage goods anticipated to workers, respectively. The first function is illustrated in Figure 2.1, where it is indicated by $Y(w_t)$.

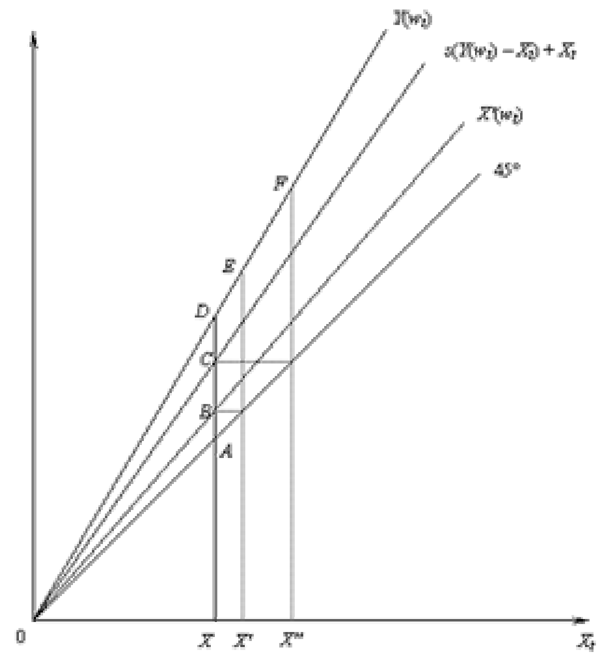


Figure 2.1 – The Smithian–Marxian model

Note that $Y(w_t)$ rotates clockwise as the wage rate increases. Suppose that the population rate of growth is $G_n(w_t)$. Since the capital/labour ratio is a constant, if the stock of capital at time t is X_t, then in order to maintain full employment at the given wage rate in the next period, it is necessary to supply $(1+ G_n(w_t))X_t$ at the end of period t as capital. This requirement is illustrated in Figure 2.1 by the curve $X'(w_t)$. This curve rotates anticlockwise as w_t increases. The vertical distance between the curve $Y(w_t)$ and the 45° curve indicates the surplus for each level of initial total capital. Since there is no rent, this surplus makes up net profits. Hence, curve $s(Y(w_t) - X_t) + X_t$, giving the amount of savings out of profits plus the total capital, indicates the capital stock that capitalists will employ at time t + 1. In the case illustrated in Figure 2.1, for example, for the level of total capital OX, the segment AB indicates the amount of net saving (i.e. net of total capital) necessary to ensure continuous full employment over time at the given wage rate (the segment XB indicates the amount of total resources which must be saved in order to ensure an increase in demand of labour equal to the increase in supply). This amount can be interpreted as the *net demand for capital* for steady-state full-employment growth. The segment *AC* indicates the amount of net savings (i.e. net of total capital) available for employment in productive activities. This magnitude can be interpreted as the *net supply of capital*. For the sake of simplicity we assume that at time *t* there is full employment of labour, i.e. $L_t = N_t$. In Figure 2.1, if total capital at time *t* is *OX*, then at time $t + 1$ the demand of labour for full employment equilibrium is given by the segment *X''F* while the supply of labour is *X'E*. Hence the supply of capital is greater than the demand of capital for steady-state full-employment, wages will increase and also the rate of population growth will increase. The former change yields a clockwise rotation of curve $Y(w_t)$ and, therefore, of curve $s(Y(w_t) - X_t) + X_t$, and an anticlockwise rotation of curve $X'(w_t)$. The equilibrium between net demand and supply of capital for steady-state equilibrium will be reached when wages reach a value w^e such that demand and supply of capital for steady state equalises:

$$(1+ G_n(w^e)) = s/(v + w^e).$$

From this intuitive description, it can be seen that for classical economists adjustment on the capital market between savings and investments is not the same as in neoclassical economics, since it occurs mainly through adjustments in the labour market. Moreover, unlike neoclassical growth theory as we shall see, the rate of growth of the economy is determined by the interplay between savings and population growth rate, the former being completely employed in investment and the latter being endogenously given as an increasing function of the real wage rate.

2.2.2. The von Neumann Model

Von Neumann studied a multisector version of the classical Smithian–Marxian model. From the formal point of view, this model is a multisector linear model with only labour as the non-produced means of production and possibly with joint production. Von Neumann looks for an activity level vector yielding the maximum rate of growth of the system and the associate competitive price system. He deals with labour in exactly the same way as classical economists did, that is, labour does not appear explicitly in his model because it is ‘produced’ by linear technology by means of wage goods (Champernowne, 1945–46; see also Kurz and Salvadori, 1993). Hence production is carried out by means of physical capital and wage good capital, and the aim of the productive activity is exclusively the accumulation of capital. The model is characterised by a zero value of rent and a set of

production techniques M_t which is stationary and is represented by a finite set of linear processes. Finally, X_t and Y_t are vectors.

The equilibrium concept considered by von Neumann is the balanced growth equilibrium in which the economy is growing at the maximum technical rate. He provides sufficient conditions ensuring the existence of a semi-positive equilibrium vector of activity levels, a semi-positive equilibrium price vector and a non-negative rate of interest which is equal to the (maximum) rate of growth.

The multisector von Neumann model can be considered the first complete economic model in which the rate of growth is endogenously determined. Some of the elements that characterize this model, as the assumption that all factors of production are producible, have been recently used in the modern endogenous growth literature to obtain persistent growth in absence of technological change (e.g. Lucas, 1988; Jones and Manuelli, 1990; Rebelo, 1991). This approach to generating endogenous growth will be discussed in Section 2.5.

2.3. THE KEYNESIAN TRADITION

2.3.1. The Harrod–Domar Model

Harrod (1939, 1948) and Domar (1946) developed the first macroeconomic model to formally analyse the problem of growth. In so doing, particular attention is paid to make explicit the relationship between the consumption-saving by households and the investment decision by entrepreneurs, although these behaviours are not theoretically developed. In fact, the consumption-saving decision is defined, following the Keynesian approach, by an exogenously given propensity to consume, while the investment decision is defined by the accelerator principle. In their model, production is obtained only by means of physical capital and labour.[8] Given the usual Keynesian assumption of fixed prices, firms choose the best technique at the given prices. Thus generically there is only one cost-minimising technique, which implies that the capital/labour ratio and the capital/production ratio are uniquely determined (By using both the normalisation and the notation of the previous section, we have $K_t/L_t = K_t/Y_t = v$). Since Harrod and Domar, following Keynes, believe that the market mechanism is not able to attain full employment of labour, they focus only on the equilibrium of the goods market – which holds when savings are equal to the desired investment – rather than the general equilibrium on the goods and labour markets.

An economy growing along a path with equilibrium on the goods market is said to be on its warranted growth path. Along this path one obtains

$G_w = s/v$ where G_w is called the *warranted* rate of growth of income, s is the rate of saving and v is the capital/production ratio. The behavioural hypothesis on producers and the Keynesian multiplier yield that the warranted growth path is unstable. If the warranted growth path ensures also the full employment of labour – a possibility which is just accidental in this model – the economy is said to be on the *golden age* growth path.

The decision structure of the model is the following. At time t the economy is endowed with a technology M_t, which is assumed to be subject to exogenous technical progress, a stock of capital X_t, and a given amount of labour N_t. Entrepreneurs decide the level of investment at time t on the basis of the acceleration principle, according to which the optimal level of investment depends upon the expected desired level of capital employed at time $t+1$, X^e_{t+1}, and the current level of capital stock. In equilibrium the former must be equal to the actual level. By the multiplier principle, the investment at time t determines the equilibrium level of production at this time. Given the production function and the hypothesis on prices, the cost minimising technique, M^*_t, is chosen and the desired amounts of capital and labour at time t are obtained, $K_t/v = L_t = Y_t$. Assumed to be on a warranted path, the current amount of capital X_t must be equal to the desired capital stock K_t. In this model there is nothing ensuring the equilibrium on the labour market, i.e. it is not necessary that $N_t = L_t$.

Aggregate income Y_t determines the aggregate amount of consumption C_t and hence the aggregate amount of saving S_t through the Keynesian assumption of constant propensity to consume; i.e. $C_t = (1-s)Y_t$ and $S_t = sY_t$. Through the equilibrium condition on the goods market one obtains that $sY_t = I_t$, i.e. the saving decision of households is coherent with the investment decision of producers.

Figure 2.2 describes the process of accumulation in the Harrod–Domar model. The level of production that ensures full employment of the capital stock is described by curve Y_t, while the saving function is function sY_t. By assuming that the decay rate of capital is δ, $0 \le \delta \le 1$, the curve δK_t indicates the amount of resources which must be devoted to replacing the decayed capital, while curve $(n+\delta)K_t$ indicates the amount of resources necessary to ensure the reintegration of capital and an increase in the demand for labour equal to the rate of growth of population. According to Harrod and Domar's view, investment determines the level of income, Y_t, which determines in turn the net savings, i.e. the net supply of capital for full employment of capital steady state, which is indicated by the segment *AC*. The current stock of capital, which is optimal by the accelerator principle, *OK*, and the rate of growth of population determine the net demand of capital for full employment of capital and labour steady state, indicated by segment *AB*. In the case illustrated in Figure 2.2, the supply of capital yields a demand for

labour at time $t+1$, *DE*, greater than the supply of labour available at that time, *FG*. Harrod and Domar, being concerned mainly with steady-state conditions, do not consider in detail what happens in this case, although it is suggested that an increase in wages and a subsequent inflationary process could be generated.

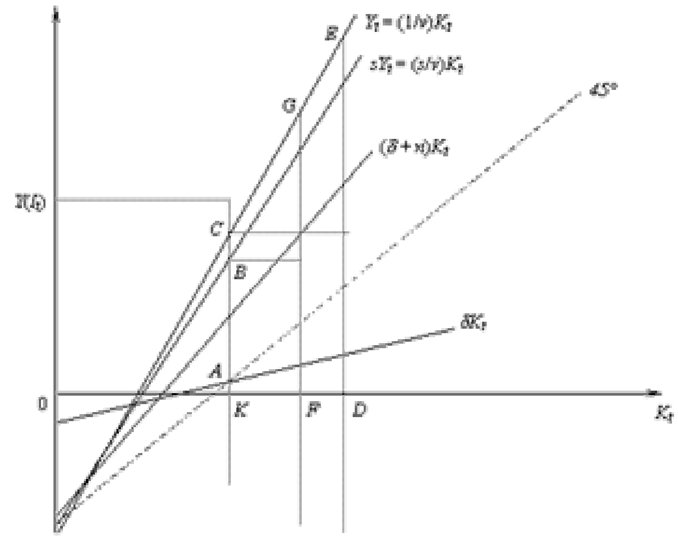


Figure 2.2 – The accumulation process in the Harrod-Domar and Kaldor models

2.3.2. The Kaldor and Pasinetti Models

Kaldor (1954a, 1954b, 1961) holds that it is not saving, investment, technical progress and population growth that are the causes of growth – these being just features of growth – but the attitude of investing by society and in particular of entrepreneurs. In this he follows the Keynesian approach in conceiving the expansion of the economy as driven by psychological and social factors like 'human attitude to risk-taking and money-making' (Kaldor, 1954a, p. 228).

In seeking a growth theory which explains the real dynamics of economies, Kaldor criticises Harrod–Domar's model on the grounds that it explains only the growth of an acyclical economy with full employment of savings rather than the *actual* rate of growth of a system that does not maintain a shifting equilibrium. Indeed, he held that in a system in which growth results from successive booms and slumps, the actual trend is determined by the 'natural rate' of growth (Kaldor, 1954a). Because of the sociological factors underlying the phenomenon of growth, he maintains, following here the fundamental Schumpeterian intuition, that a satisfactory growth theory cannot be constructed without a business cycle theory (Kaldor, 1954a, 1954b, 1957). However, he never formally develops his position on economic growth, and his major contribution consists in solving in an original way the stability problem of the Harrod–Domar model. This is accomplished by allowing the possibility that the economy can grow along a natural growth path through adjustments in the rate of saving due to changes in the distributive shares between wages and profits and assuming that the population rate of growth is constant. The latter assumption is adopted only in his formal works, while in several non-technical articles (Kaldor, 1960) he and other post-Keynesian economists like Joan Robinson (Robinson, 1963) share the classical view that the rate of growth of population depends upon the wage rate and, therefore, upon the rate of growth of the economy.

Kaldor's contribution can be seen in Figure 2.2, where the usual production function (Y_t), supply of capital (sY_t), demand of capital for steady state with full employment ($(n+\delta)K_t$) and demand of capital for reintegration (δK_t) and demand for labour curve (L_t) are depicted (we still assume that the there is only one method of production). Note that now the average saving rate depends inversely upon the wage rate due to the relation

$$s = s(w_t) = s_w Y_t + (s_p - s_w)\Pi_t(w_t),$$

where s_p is the saving rate out of profits, s_w the saving rate out of wages and $\Pi_t(w_t)$ the profits (Kaldor, 1955–56), which are inversely related to the wages. If the stock of capital is *OK*, then the supply of capital, *KC*, is higher than the demand of capital ensuring full employment and constant capital/labour ratio, *KB*. At time $t+1$ the demand of labour, *DE*, is higher than the supply of labour, *FG*, therefore wages increase. This yields an increase in s and a clockwise rotation of the curve $s(w_t)Y_t$, and L_t tends to N_t. The equilibrium is attained when w_t is such that $s(w_t)Y_t = (n+\delta)K_t$; in this case, the supply and the demand of capital are equalised at the level which ensures full employment. Although the equilibrating mechanism is described in terms of forces operating in the labour market, Kaldor (1961, pp. 196–7) suggested that the process could equally take place in the goods market through changes in prices.

Kaldor's approach has been developed and pursued further by Pasinetti (1962) and by a fairly vast literature which, however, in emphasising the study of equilibrium paths have shifted attention from the original attempt to construct a growth theory out of a business cycle theory to the more traditional view of constructing a theory of economic growth at the natural rate.

2.4. NEOCLASSICAL GROWTH MODELS

2.4.1. The Solow Model

A different attempt to solve the stability problem of full-employment steady state is that of Solow (1956). He accomplishes this task by assuming a neoclassical production function that allows for flexible coefficients of production. By adopting a neoclassical framework, Solow changes the object of analysis with respect to the Keynesian growth theory; in his view the major problem is to construct a theory of general full-employment growth, and the most important concern is to ensure the convergence of the economy towards the natural growth path. Hence, growth theory has to explain the *potential* growth of economies (Solow, 1999, 2000), without paying attention, therefore, to cyclical trends of the economy and their possible effects on the long-run trend of the economy.

Solow assumes that there is only capital and labour as factors of production. The technology is represented by means of a neoclassical production function with constant returns to scale, decreasing productivity with respect to physical capital and possibly labour-augmenting technical progress. In order to construct a model that conciliates full employment of resources with growth, Solow assumes that prices are flexible and therefore all markets are cleared. In particular, the equilibrium on the capital market yields that investments are equal to savings while the equilibrium on the labour market yields that there is always full employment of labour.

Production is distributed between savings and consumption on the basis of a Keynesian saving rule. If savings are equal to the level of investment which ensures the constancy of per capita capital with full employment, then the economy is in the steady state. Otherwise price adjustment on the capital market yields equalisation between savings and investments and appropriate changes of the per capita capital until the steady state is attained. The convergence process towards the steady state is ensured by the assumption of decreasing productivity of capital.

At time t the economy is endowed with a technology set, a given amount of capital and labour. Producers choose the best technique according to the ruling prices, and full employment of resources is ensured by their complete flexibility. The resulting income is either consumed or saved according to the usual, exogenously given, propensity to save, s. Savings are completely transformed into investments by the flexibility of the interest rate on capital market. Investment and the initial stock of capital determine the amount of capital available at time $t+1$. The technology set and the amount of labour available at this time are determined exogenously by the values of these

variables at time t, the former by assuming a positive rate of technical progress, the latter by a positive rate of population growth.
Figure 2.3 illustrates the accumulation process in the Solow model. In this figure, all variables are expressed in per capita terms. Curve $(n+\delta)k$ indicates the (per capita) demand of capital for a full-employment steady state with constant per capita capital, while curve $sf(k)$ indicates the (per capita) supply of capital. If the capital per capita is *Ok*', the supply of per capita capital, *k*'*A*, exceeds the per capita demand of capital for a full-employment steady state with constant per capita capital, *k*'*B*. Hence, in order to maintain full employment on the labour market the rate of interest will decrease and entrepreneurs find it profitable to increase the demand for capital. The price adjustment and the consequent demand increase is such as to equalise demand with supply. At the new, higher level of capital per capita production will generate a new supply of capital and a new demand of capital for a full-employment steady state with constant per capita capital will arise. If the former is equal to the latter, a steady state is attained, if the former is still greater than the latter a further decrease in the interest rate and of the per capita capital occurs. And so on.

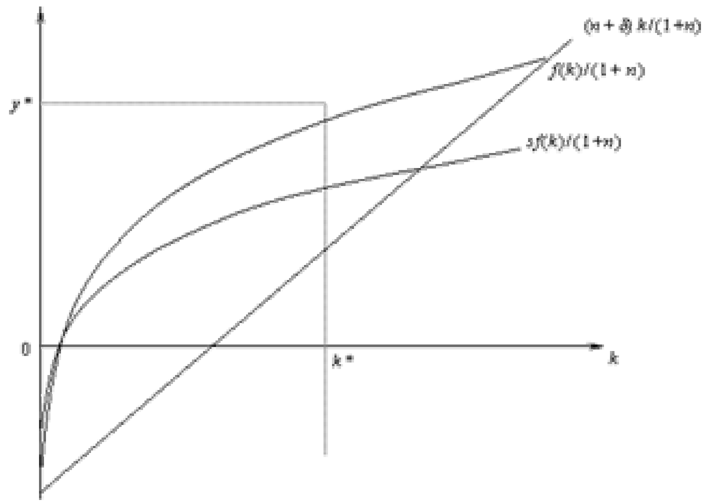


Figure 2.3 – The Solow model in discrete time and without technical progress

2.4.2. The Growth Model à la Ramsey

Inspired by the article by Ramsey (1928), several growth models have been constructed in order to improve Solow's model by making the rate of

saving of households endogenous. The early models following this approach are characterised by a *normative* interpretation of the accumulation process in that the economic decisions concerning production and saving are taken by a planner choosing over an infinite horizon. An important peculiarity of these early contributions is that the production side of the economy and the income employment side are not separated since a single agent, the planner, has control both over production and saving decisions. The literature later emphasised that the optimal path chosen by the planner coincides with the path chosen by a perfectly competitive economy with many agents. This step is important in order to ensure that Ramsey's approach is able to provide a *positive* theory of growth like the traditional models, such as Harrod–Domar's and Solow's. The decentralisation problem, moreover, opened the way to the construction of growth models in which the equilibrium path is analysed independently of the optimal, centralised path. This frees the theory from the hypothesis of perfectly competitive markets or from the assumption of the absence of external effects, and it is one of the most important contributions of the recent endogenous growth theory.

In the traditional growth models *à la* Ramsey (Cass, 1965; Koopmans, 1965), the assumptions concerning the production function are the usual neoclassical ones and the planner is endowed with a separable and stationary utility function $u(\cdot)$ and a constant discount rate ρ. The optimal path, which dynamically has the structure of a saddle point, is unique and converges towards the steady-state path.

The basic structure of the normative version of Ramsey's model can be illustrated as follows. At time t the economy is endowed with a technology set M_t and given amounts of capital X_t and labour N_t. The planner chooses the best method of production M_t^* and the entire amount of labour and capital to produce the output, Y_t. The planner then decides how to allocate the amount Y_t between consumption and savings which are immediately transformed into investment, I_t. The stream of consumption levels $(C_t)_{t=0}^{\infty}$ is chosen in such a way as to maximise the sum $\sum_0^\infty \rho^t u(C_t)$ subject to the constraint $X_{t+1} = Y_t - (1-\delta)X_t - C_t$, where δ is the capital decay rate.[9] Investment at time t and the initial stock of capital X_t determine the stock of capital available at time $t+1$, X_{t+1}. The amount of labour available at time $t+1$, N_{t+1}, is determined by a demographic rule that is given exogenously.

The accumulation process in Ramsey's model is similar to that of Solow's described by Figure 2.3. Now, however, along the optimal path the rate of saving changes over time and converges towards the long-run level associated with the steady state.

2.5. ENDOGENOUS GROWTH MODELS

The aim of the endogenous growth theory is twofold: first, to overcome the shortcomings of the Solow and Ramsey models which are unable to explain sustained growth, and second, to provide a rigorous model in which all variables which are crucial for growth, in particular savings, investment, and technical knowledge, are the outcome of rational decisions. Hence the endogenous growth theory has adopted Ramsey's model as a reference theoretical structure, in which saving is the outcome of a maximising agent and the equilibrium growth path is seen as the consumption/saving trajectory chosen by rational agents by solving an intertemporal optimisation problem.

As for the former, the endogenous growth literature points out that a necessary condition for perpetual growth is that from the household's point of view the rate of interest should never be driven too low, and this is ensured if the productivity of accumulated factors does not decrease to zero as accumulation proceeds (see for example, Jones and Manuelli, 1997). On the contrary, if this case occurs, savings will be driven to a level that is not enough for fuelling sustained growth. In this perspective, the main object of the endogenous growth theory has been to develop economically meaningful ways of ensuring non-decreasing returns to scale with respect to the accumulated factors. This has been accomplished either by removing the scarcity of natural resources or by introducing technical progress. As far as the former is concerned, for example, labour has been straightforwardly transformed into a fully reproducible resource, human capital. As for technical progress, one of the main features of the endogenous growth theory is the capacity of endogenise the investment decision yielding technological progress which consists mainly in the introduction of new intermediate and/or final goods (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992, 1998).

2.5.1. Accumulation of Physical and Human Capital

The simplest endogenous growth model is the so-called AK model (see e.g. Rebelo, 1991). It can be considered as a Ramsey model with the assumption that the production function is linear with respect to physical capital and that scarce resources are not considered explicitly. Within the AK model it is possible to obtain a constant positive rate of growth of the per capita consumption along the optimal path. This path, moreover, can be rationalised as the outcome of a decentralised perfectly competitive economy.

A second approach to obtain sustained growth is to introduce human capital (see e.g. Lucas, 1988). By following this approach, persistent growth is obtained by transforming labour from a scarce resource into a fully reproducible factor by interpreting it as human capital. However, now an additional decision has been taken by households concerning the amount of resources to employ in the accumulation of physical and human capital.

The following is a summary of the structure of the Lucas model. At time t the economy is endowed with a technical set for producing output M_t, a technical set for the production of human capital MH, given stocks of physical capital X_t and human capital H_t. Set MH is made up by only one linear method of production and does not change over time. On the basis of the ruling prices, households choose the amount of human capital to employ in productive activities, uH_t, and in the further accumulation of human capital $(1 - u)H_t$. Firms, on the other hand, choose the best technique M_t^*, the amount of capital K_t and the amount of human capital to employ in production. Flexibility of prices ensures full employment of factors. Households allocate the production Y_t between consumption and savings according to their utility function, and savings are, as usual, transformed into investment through the capital market. At time $t + 1$ the economy will be endowed with a technical set M_{t+1}, which is different from that at time t because of exogenous technical change and externalities from the accumulation of human capital, a technical set MH, capital stock X_{t+1} determined by initial capital stock X_t and investment I_t, and, finally, a stock of human capital H_{t+1} determined by investment $(1 - u)H_t$ and the production process in MH.

The accumulation process in the endogenous models considered here is similar to that illustrated in Figure 2.2 for the Harrod–Domar model. Similar to this model, along the optimal path the rate of saving is constant; however, unlike the model, it is determined endogenously by the maximizing behaviour of the planner or of households. Moreover, the AK model, unlike the Keynesian models, follows the classical and neoclassical tradition in conceiving investment as induced by savings, rather than the other way round.

2.5.2. Technical Progress

In the Lucas model externalities are present but they are not essential to ensure sustained growth. An alternative approach to the accumulation of factors to ensure sustained growth is the introduction of technical progress. Technical progress can be considered an improvement in technological knowledge incorporated in the new production function – usually due to unintentional effects like externalities.[10] This class of models is important in the development of a growth theory since, because of externalities or explicit introduction of imperfect competitive markets, it makes a substantial break with respect to the growth models *à la* Ramsey given that on the equilibrium path there is no longer an optimal allocation of resources. This implies that analysis of the optimal path can no longer be obtained by means of the normative approach but has to be carried out through a decentralised mechanism which requires a detailed description of the behaviour of agents and of the working of market mechanisms.

The structure of the model with externalities is very similar to that of Ramsey's model, except that the existence is assumed of a finite number of firms whose production and accumulation decision affects positively and unintentionally the technology of all firms. This effect ensures the constancy of the productivity of the accumulated factors, thereby removing the tendency towards a growth path in which the rate of growth is determined only by exogenous factors like population growth.

Models with new products are usually three-sector models: a final good sector, an intermediate sector and an invention sector which produces 'designs' of new intermediate products. Competition is perfect in the final and invention sectors, while that of intermediate goods is imperfectly competitive because there is a fixed cost associated to the purchase of new designs. Imperfect competition in the intermediate sector is necessary to ensure the existence of profit in this sector which, through fixed costs, will be transferred to the invention sector. These profits in turn are the incentive for the invention sector to continue its activity. Such models endogenise the production of new ideas and introduce the Schumpeterian idea that technical progress is linked with imperfect competition. However, it is not possible to consider this class of model a truly Schumpeterian one since the production of new ideas is here conceived as a smooth process, while in Schumpeter invention is strictly linked, as in Kaldor's view, to business cycle.

In the basic model taken from Romer (1990), there exists only one final good, which can be employed also as physical capital, infinite intermediate goods and potentially producible 'ideas'. Each intermediate good is produced by means of physical capital and only one idea. The technology set for producing the consumption good is assumed to be stationary, while it is assumed that the production sets for producing ideas and intermediate goods are made up by only one method of production. The method of production to produce ideas is linear and requires only labour, while the method of production of intermediate goods requires physical capital and ideas. It is linear with respect to physical capital, although it has fixed costs due to the use of the corresponding idea. At time t the economy is endowed with technology sets to produce new 'ideas', intermediate goods and the consumption good indicated respectively by MI_t, MIN_t and MC, a given stock of physical capital X_t and a given amount of labour N_t. Households choose how much labour to allocate to produce consumption goods and to produce ideas. The amount of research labour, together with the technology for

producing ideas, obtains new ideas affecting the technology set for producing intermediate goods available in the next period. Intermediate goods are produced by employing the whole physical capital and the relevant technology. Intermediate goods, in turn, together with labour for the production of final goods and the relevant technology, produce the final good. This is allocated by households into savings and consumption. The former, through the capital market, are transformed completely into investment, which determines the stock of physical capital available in the next period, X_{t+1}. The production set for producing the final goods at time t is assumed to be invariant over time, while the set for producing ideas at time $t+1$, MI_{t+1}, is different from that at time t because of externalities from the existing amount of ideas. The latter is interpreted as learning by doing or as the public nature of researchers' skills. Unlike externalities in Lucas' model, here the existence of externalities is important in guaranteeing sustained growth since it ensures the linearity of the production function in research for new ideas.

2.6. CONCLUDING REMARKS

In this paper we surveyed in a simple and intuitive way different models for analysing the accumulation process conceived in various traditions from classical economics to the endogenous growth theory. Our main aim was to highlight similarities and differences among these alternative theories. We pointed out that there is continuity from classical to endogenous growth theory, partly through Keynesian theory (see also Hahn and Matthews, 1965, pp. 8–9), concerning the fact that the steady state is conceived as endogenously determined by the model. By contrast, neoclassical economists see it as exogenously determined by factors considered outside the realm of economic explanation. We also emphasised continuity between classical, neoclassical and endogenous growth theory, as opposed to Keynesian theory, in terms of the saving/investment relationship. While the former theories conceive saving as wholly transformed into investment, and therefore, growth being determined by saving itself, Keynesian theory conceived investment as the source of growth and no relationship between the former and the latter variable necessarily exists.

NOTES

* We would like to thank Heinz Kurz for valuable comments.

1. Smith, for example, while allowing that wages and rents could be employed in savings, maintains that capitalists are 'naturally the most disposed to accumulate' (WN IV.viii.61).

2. For a more complete elaboration on this point see Rosenberg (1975).
3. See also Hollander (1973, p. 158) and Eltis (1984, p. 88).
4. This assumption does not imply that there is full employment of labour. An ‘industrial reserve army’ is compatible with the case considered.
5. It seems reasonable to say that Marx considered explicitly investment to improve technology. In the basic classical model here considered we shall neglect this fact and we consider technical improvement due to non-deliberate economic decisions, like in Smith’s view (see McNulty, 1968). Investment in R&D will be taken up in detail later, in dealing with the endogenous growth theory.
6. Smith holds that labour and capital have an increasing productivity (WN I.iii, I.viii.57, II) and Marx seems to hold that labour and capital have constant productivity (Marx, 1956).
7. Figure 2.2 does not appear to reflect some important features of the Marxian view of the accumulation process; in particular, the full employment condition. However, this figure is able to make the classical approach comparable to the more recent models as far as the adjustments on the capital market is concerned.
8. Harrod’s model is quite different from Domar’s. In what follow we shall refer to *the* Harrod–Domar model because we consider only a simplified version of these models.
9. In the original version of the Ramsey (1928) model it was assumed that $\rho = 1$. The fact that future utilities were not discounted implied that many feasible paths had an infinite value. To compare these paths, Ramsey recast the maximisation problem as a minimisation problem in which the objective was the sum of losses from a reference path (the Bliss point).
10. See e.g. Romer (1986), or the introduction of new products that can be either intermediate goods (Romer, 1990; Aghion and Howitt, 1992) or final goods (Grossman and Helpman, 1991).

3. Endogenous growth theory as a Lakatosian case study

Mario Pomini

3.1. INTRODUCTION: THE ISSUE

In the mid 1980s, growth theory experienced a remarkable revival and became once more a very active area of macroeconomic research. Starting from the seminal articles by Romer (1986) and Lucas (1988), the research took a precise direction in the sense that, in contrast to the earlier neo-classical view, it called for an endogenous determination of technological change, which means an endogenous determination of the sources of growth.

This article will concentrate on the treatment of endogenous growth by neo-classical growth theorists. It uses the Methodology of Scientific Research Programmes (MSRP) proposed by Lakatos (1970) in order to explain why the endogenous growth approach was not incorporated into the neo-classical growth programme until the late 1980s, although the essential features were well known during the 1960s. The resulting thesis is that the new growth theory may be seen in terms of an extension of the neo-classical research programme to incorporate theoretical elements which previously fell beyond its scope.

MSRP was introduced into economics methodology by Latsis (1976) and in the following decades was used intensively as a way to study the nature of progress in many branches of economics.[1] Even if the MSRP is not without its critics among economic methodologists (for example Hands, 1990), it remains a useful framework within which to analyse the evolution of economic ideas.

3.2. THE DEVELOPMENT OF THE NEO-CLASSICAL GROWTH RESEARCH PROGRAM

In this section and in the following, we will try to characterize neo-classical growth theory in Lakatosian terms. The important organizing category in

Lakatos' methodology is the *scientific research programme*, which is a more detailed version of Kuhn's concept of paradigm. Specifically, it should be thought of a series of theories that make up a continuous whole because they share some common elements. A research programme is made up essentially of two elements: a hard core with its protective belt and its positive heuristics. The core hypotheses are viewed as unchangeable by those who take part in the research programme and anomalies are incorporated into the programme not by changing the indispensable hypotheses but rather by modification of the auxiliary hypotheses, the initial conditions and the observation set.

Further, each research programme is characterised by a set of methodological rules: some of these (the positive heuristics) indicate what the researcher should do, whereas others (the negative heuristics) are injunctions concerning what not to do. The negative heuristic has a purely protective role and aims to prevent the hard core's propositions being changed and tested. The positive heuristic outlines development directions of the research programme and consists of a set of suggestions concerning how to change or modify some aspects of the research programme which may be refuted (Lakatos, 1978, p. 110).

A research programme is not a static entity but evolves over time: new facts are discovered and new problems arise which require changes in the protective belt and the positive heuristic. It may be *progressive* or *degenerating*. If a research programme leads to the discovery of new facts which are successfully explained on the basis of the programme elements, then there is a *progressive problem shift*. If, on the other hand, the programme is running out of steam and the new hypotheses added are ad hoc hypotheses, and thus partial adjustments which do not increase the empirical content of the programme, then the programme undergoes a *degenerating problem shift* (Lakatos, 1978, p. 118).

To illustrate the essential features of the neo-classical research programme it will be easier to divide the exposition into two parts. In this section, we will consider the short-run dynamic model. In the next section, we shall see under which hypotheses the model can be extended to encompass long run phenomena and obtain a theory of economic growth.

The best illustration of the guidelines and the most important results of the neoclassical growth programme remains Solow's essay, *Growth Theory: An Exposition* (1970). From this it can be gleaned how the neo-classical hard-core[2] organised around the following propositions:

HC1) Growth theory concerns itself with the conditions under which an economy grows in steady-state conditions.

HC2) The dynamics of an economic system is determined by the accumulation of the factors of production.
HC3) The supply side of the economy is described by an aggregate production function which allows complete substitutability of factors of production, which typically are labour and capital.
HC4) The aggregate production function is characterised by constant returns to scale on the factors employed (the production function is homogeneous of degree one).

These propositions are not the same in nature. The first two are wholly general conditions which can be found also in research programmes other than neo-classical ones. HC1) was introduced by Harrod (1939) who was the first to pose the problem of formalising the way in which an economic system can reach a position of long run steady growth. For Harrod, the idea of steady-state growth is nothing other than the adaptation to the dynamic case of the notion of equilibrium which we find in statics, even if this move represented 'a real revolution in mentality' (Harrod, 1939, p. 15). Proposition HC2) goes even further back in the history of economic thought and translates the idea of the classical authors according to which economic growth appears essentially as a circular process: in order for the economic system to expand, part of the output must be saved and used to increase the stock of factors of production, primarily, capital. Economic growth and factor accumulation are two processes which, from Adam Smith on, largely coincide.

Propositions HC3) and HC4) are typical aspects of the neo-classical approach. HC3) makes it possible to endogenise the capital–labour ratio, which in Harrod was constant. Solow's criticism of Harrod is that the latter chose to study long-run phenomena using short-run technical tools, such as that which requires technical coefficients of production to be constant. HC4) comes from the theory of distribution and requires that output be shared among the factors according to their marginal productivity. These two propositions, on the basis of which the capital–labour ratio may vary and its variation is regulated by the quantity of available capital for the economy, enabled the young Solow to provide a brilliant solution to the problem left unsolved by Harrod of instability on the economic growth path.

The positive heuristic of the neo-classical research programme can be illustrated by the following propositions:

PH1) Build models where the agents optimise.
PH2) Build models which allow predictions on the equilibrium states.
PH3) Show up the logical supremacy of resource allocation over distribution.

PH4) Do models where the long run dynamics is determined by the accumulation of physical capital.

The first two propositions are typical of the neo-classical view, in the broadest sense, and do not require any particular attention. EP3) expresses the fact that dynamic allocation of resources is determined purely by technology and by the initial stock of resources. Solow here shifts radically from previous tradition which considered growth and supply as two fundamentally inter-dependent processes and profit was the driving force in economic growth. EP4) is representative of the deeply-held conviction, shared by those studying economic growth in the '50s and '60s, that economic development was driven by the accumulation of capital and so by industrialisation. This does not mean that the importance of other factors was not recognised, but simply that they did not have any great weight in theoretical elaboration (Arndt, 1984).

3.3. FROM THE DYNAMIC MODEL TO THE THEORY OF GROWTH

It is Solow himself (1970, p. 34) who observed that the solution to the problem of Harrodian instability, albeit an undoubted step ahead on the theoretical plane, does not provide a solution to Smith's problem: to identify which factors determine growth in the long run.

To move from the dynamic model to growth theory, as required by empirical evidence, we must add a further element to the model, so far missing, which can support the dynamics over the long term. With his usual clarity, Solow states that 'there are two obvious candidates: technological progress and increasing returns' (1970, p. 34). Both exogenous technological progress and increasing returns to scale are tools which can offset the consequences resulting from diminishing returns in the accumulated factor. Exogenous technological progress obtains this effect via continuous innovation which postpones the production function and returns to scale via those processes of cumulative causation which are linked to the increase in the size of the economy. In both cases, the result is an increase in per capita output and capital, with a single restriction that, to be compatible with the steady-state condition, both returns to scale and technological progress must be introduced into the model in a very particular way.

As is well-known, Solow chose exogenous technological progress. So the list of hypotheses included in the hard-core must be supplemented by:

HC5) Technological progress tends to be a temporal exogenous trend which increases output at a constant rate.

According to HC5), in the long run, the growth rate of the economy coincides with the growth rate of labour efficiency, to which must be added, if necessary, the growth rate of the working population. The continuing increase in labour efficiency creates new opportunities for profit which, in equilibrium, exactly compensate its downward tendency following the process of accumulation.

From a methodological point of view it is not difficult to see how HC5) is a rather typical situation in which a theoretical model is adjusted to empirical reality simply by the introduction of an ad hoc hypothesis. In the methodological literature, the notion of ad hoc hypothesis refers to the various stratagems used by researchers to introduce new assumptions solely to save the model in the face of contrary empirical evidence. An hypothesis is ad hoc when it cannot independently produce new predictions and hence does not improve the empirical content of a theory. Following Popper, we can say that the adjustment of a theory for the sole purpose of protecting it in the light of contrary evidence is not good scientific practice. The modification introduced into the dynamic model with the addition of exogenous technological progress was ad hoc because it did not lead to any further predictions, backed up in turn by further checks and observations. The empirical content of the theory was reduced rather than increased.

But there is also a further aspect beyond the typically Popperian, which characterises an hypothesis as ad hoc and this aspect has been highlighted by economists. Lakatos claims that an hypothesis may be ad hoc not just because it prevents the genuine falsification of a theory, but because it conflicts with the programme's heuristic and so weakens its internal coherence and unity. This is the case for theorists of rational expectations (Hands, 1988) who hold that other types of expectations have been damaged by ad hoc assumptions, not so much because they are contrary to available evidence but rather that they are inconsistent with the principle of maximisation which represents the key assumption of neo-classical economics. The same criticism can be levelled at the hypothesis of exogenous technological progress: it is an ad hoc hypothesis not because it lacks the capacity to produce new empirical evidence but because it is not derived from the optimising behaviour of the individual agent and therefore it is inconsistent in microeconomic terms.

It is a remarkable instance in the history of economic thought that a research programme with such a flimsy empirical basis and supported by an ad hoc assumption should have gained such an important position as to hold the stage for two decades before being challenged by recent models. Once again it is Solow who offers us the key to understanding the success of his approach to economic growth:

> I shall concentrate on technological progress without taking returns to scale into account, for two reasons. In the first place, I reckon that technological progress must be the more important of the two [elements under consideration] in real economy. It is difficult to believe that the US is enabled to increase output per man at something over 2 per cent a year by virtue of unexploited economies of scale. ... Second, it is possible to give theoretical reasons why technological progress might be forced to assume a particular form required for the existence of a steady state. They are excessively fancy reasons, not altogether believable. But that is more of a lead than we have on the side of increasing returns (Solow, 1970, p. 43).

Here Solow clearly puts forward two different arguments, equally decisive for our ends. The first is a meta-analytical argument which can be defined as rhetorical. The second is the formal argument that shows how a mathematical structure may be found which is coherent with the basic viewpoint.

The argument whereby the technological progress hypothesis seems to Solow to have greater persuasive force must be qualified in that Solow is here referring not so much to growth theory as to the whole structure of neo-classical economics. What appeared convincing was the fact that the Solow parabola may be considered the extension to the dynamic case of the theory of general economic equilibrium with all the supports of associated hypotheses, including the essential one regarding constant returns to scale. If the essential aim of growth theory was to offer a dynamic vision of the Arrow–Debreu model, it follows naturally that the principle of increasing returns was almost completely abandoned as it was incompatible with the theory of competitive equilibrium.[3]

So it was not empirical evidence or analysis of facts which indicated the research path to be followed but rather a theoretical vision in which the economy tends naturally towards co-ordination of economic decisions through the price mechanism and is highly suspicious of intervention in the economy. All that could challenge such a perspective was set aside by methodological decision. This rhetorical vision which aimed to justify a particular vision of the economic system hinging on free-play in the market was not without consequence for the development of the research programme. Neo-classical theory was gradually consigned to irrelevance on the empirical plane and its position became even weaker when it was strongly criticised and its theoretical foundation and internal coherence challenged (Harris, 1980). On the other hand, the theory of economic development as a separate discipline from growth theory was given a boost and great importance was held in it by all those elements which were ignored by the rival model, starting with the active role of the state (Arndt, 1984).

The second argument advanced by Solow was the necessary conclusion of the preceding one. The fact that technological progress is exogenous and in particular takes the form of increased labour efficiency, makes it possible to maintain all the model’s implications, from monotonous convergence to stability. As in the case of static equilibrium, also in dynamics the desired results in terms of existence, uniqueness and stability of the equilibrium solution are obtained.

3.4. DEVELOPMENT OF THE NEO-CLASSICAL RESEARCH PROGRAMME

For Lakatos a research programme is not a static entity. New facts are discovered, new problems emerge and consequently the protective belt undergoes some adjustment. Lakatos holds that a research programme should be evaluated on the basis of its ability to evolve with time.

In the 1960s the neo-classical research programme was enjoying most favour; new hypotheses were added to the base model which seemed to be able to increase its ability to interpret and its analytical richness via subsequent progressive shifts. In economics, a useful criterion among the many available to establish whether a theory or a research programme is progressive consists in evaluating the relationship which exists at any moment between exogenous variables, that is, the inevitably arbitrary starting point of any research, and the endogenous variables, that is, the elements which the model sets out to explain. Generally speaking, we have a moment of progress when the programme increases its scope and variables which were originally considered exogenous are in a subsequent phase endogenised.

Seen from this standpoint, most of the theoretical research in the golden age of the neo-classical programme concentrated on the attempt to make savings and technological progress endogenous. Endogenisation of savings was the work of a group of young economists of an analytical orientation, among whom are Cass (1965), Uzawa (1965), Shell (1969), Arrow and Kurz (1970), Ryder (1967) who, in a relatively short time, successfully managed to apply the techniques of dynamic optimisation. Progress was made at the level of mathematical formalisation but there was no equivalent advance in the analysis of the factors which determine economic growth, hence at an interpretative level. Retrospectively, Shell noted that ‘the success of the Hamiltonian view in the analysis of economic growth has been rather limited’ (Shell, 1987) since, even in the refined maximising version, economic growth was completely independent of savings, or rather, of typical consumer preferences.

Let us now examine technological progress. As in the neo-classical view, but not only, technological progress is the fundamental factor that determines economic growth, it comes as no surprise that the analysis of technological progress was the main area of research in growth modelling. In the literature of the 1960s, there is a clear awareness that technological progress was to a great extent an endogenous element but the question set was how to translate this common conviction onto the analytical level.

The endogenisation of technological progress followed basically two paths. The first came directly from Kaldor's observation (Kaldor, 1957) that the means via which technological progress is realised is capital accumulation. Kaldor's idea that technological progress is incorporated in new capital goods underpinned the *vintage approach* in which capital was disaggregated by year of output. Within the neo-classical system this attempt to link investment and technological progress took hold with an essay by Solow in which he developed the assumption that 'technical innovation influences technological progress only if incorporated into new capital goods or via substitution of outdated equipment with the latest models' (Solow, 1960, p. 200). Models with capital divided by year sparked off a lively body of research (Wan, 1970) and constituted the dominant approach to the analysis of the relationship between technological progress and economic growth.

However, notwithstanding the various and fanciful ways of treating capital goods by years and the efforts made in this direction, in the long run the growth rate ended up being constant over time and equal to the rate of productivity increase of the new investment. As Phelps observed, even according to the new view the 'rate of development over the long-term depends on the rate of technological progress, not on the type of progress' (Phelps, 1962, p. 256). In the models with capital incorporated, technological progress still ends up exogenous and in a steady-state. While the year-by-year approach had the merit of making the theory more realistic, it did not change the underlying view of economic growth.

3.5. THE *OLD* THEORY OF ENDOGENOUS GROWTH

The second road followed to endogenise technological progress focused on a closer consideration of knowledge as a factor of production. This was a minor research direction but which fed an original research tradition. If acquired knowledge becomes an essential factor of production, then its growth, and so technological progress, becomes dependent on the amount of economic resources assigned to it and also on the way in which innovation spreads within the economic system. Economic growth backed up by

research and innovation becomes an endogenous process. Within this approach we can distinguish three separate research paths, each of which has focused in different ways on knowledge as a factor which can be produced, owned and accumulated.[4]

The first direction of research started with Arrow's famous article in 1962 which was the fundamental contribution of the neo-classical approach to endogenous growth up to present-day models. The innovation Arrow introduced was to hypothesise that labour efficiency was an increasing function of work experience. As each work activity implies learning, workers over time become more productive. In Arrow's model, work experience is represented by the cumulative flow of new investment as it is in the production of new capital goods that learning is shown. Arrow's original idea was that learning-by-doing was a non-linear function of total investment.

Arrow's model represents the most important attempt to propose an endogenous mechanism for growth in the neo-classical school. Although it undoubtedly marked a step forward by introducing greater realism, it has two basic limitations which have reduced its theoretical use. In the first place, long-run growth is only partially endogenous due to the learning effect but depends solely on working population dynamics as in Solow's version. In the second place, its interpretative capacity is rather weak. The effect of learning is not determined by an economic choice but is the involuntary by-product of the process of accumulation. With these two elements in mind, it is hardly surprising that Arrow's approach has been considered an ingenious sophistication of the neo-classical model but has not brought any substantial changes to the prevailing paradigm.

A second approach to endogenous growth was formulated by Uzawa (1965). The novelty of Uzawa's model lies in the fact that he explicitly inserts, alongside the sector which produces the final good, a second sector which is to produce new knowledge. Uzawa hypothesises that the efficiency of this sector which produces new ideas is a concave increasing function of the quota of labour employed in this sector. Thus the parameter which measures work efficiency in the final goods sector, the usual parameter A, does not develop exogenously. The research sector employs labour to produce new ideas which shift the production function of the final sector upwards. How strong this form of technological progress will be depends on the proportion of the labour force employed in the research and education sector. As the research sector requires only employment of labour, an economy will grow at a higher rate if a higher proportion of the labour force is employed in the sector. Uzawa's is the first model we have of endogenous growth in the modern sense, in which the driving sector is research and accumulation of knowledge plays an essential role in long-run dynamics.

The third major contribution to the first wave of studies on endogenous growth was that of Shell. His approach is particularly relevant and is expounded in a series of works which indicate a precise line of research (Shell, 1966, 1967, 1976). The analytical viewpoint Shell adopts to endogenise economic growth is very close to Uzawa's in the sense that accumulation of knowledge, hence technological progress, is made to depend on the resources allocated to the research sector. The greater the resources employed in the search for new ideas, the greater economic growth will be. However, what distinguishes the two models, and this is where Shell's originality lies, is the assumption that the new ideas may be considered a factor of production in their own right. Shell develops an original form of the function of production in which stock of knowledge is considered as a third factor of production alongside labour and capital, $Y = (F, K, A)$. Further, Shell assumes that technology can be considered a public good, and this assumption will play a central role also in the endogenous growth theory of the 1980s. Shell argues: 'Technical knowledge can be employed by any economic agent without altering either its quantity or its quality. Thus, we must think of technical knowledge as a public good – primarily a public good in production' (1974, p. 79).

If new ideas and patents are a non rival input, this requires that the hypothesis of returns to scale be abandoned and the aggregate production function show increasing returns to scale. Because of the presence of increasing returns it can no longer be hypothesised that firms operate in perfect competition, for in this case the firms would take a loss. To obviate this problem, Shell adopts, like Arrow, the Marshallian hypothesis that returns to scale are external to the firm and under this condition there exists a unique and stable state equilibrium for capital (K) and technology (A).

Interestingly, also in Shell, as already in Arrow, increasing returns are not enough to guarantee a positive growth rate in the long term without exogenous technological progress. This happens because in both models returns to scale are increasing on the production function but decreasing on the real factor which is accumulated capital. This was a well-known result in the literature of the 1960s. With increasing returns to scale, the economy tends to experience explosive growth and thus it is impossible to find a stable equilibrium, but if, at the same time, marginal productivity of the accumulable factor is decreasing, then the system reaches equilibrium. But it is precisely this condition, which guarantees the steady state properties of the model, that prevents the system from showing endogenous growth.

3.6. ELEMENTS OF THE NEW THEORY OF ENDOGENOUS GROWTH

The 1970s was a very difficult period for the neo-classical research programme. The attempt to make the theoretical apparatus more realistic led to an increase in its technical complexity, such as in the multi-sectoral models (Burmeister, 1980), which produced the opposite result. The core of the programme was no longer able to respond satisfactorily to theoretical and logical criticisms, starting with the adequacy of the notion of production function and capital as a factor of production (Harris, 1980). In addition, there was an external element: in the '70s it is macroeconomic theory which abandons the study of long-run phenomena to concentrate on the short run such as the theory of the economic cycle, stagflation, the impact of rational expectations and, more generally, the crisis of Keynesian orthodoxy.

The downturn in the neo-classical programme's fortunes came with the International Economic Association Congress in Jerusalem in 1973 which was entirely given over to growth theory. As documented in Mirrlees' introduction to a selection of the papers published the following year, this was a momentous occasion. Everyone who was actively involved in growth theory took part, with Solow in the chair, and the range of discussion can be understood from the numerous sections into which the work was divided. One section was dedicated to *Growth and Technology* and hinged on two papers, one by Shell and another by Weizsacker, both dealing with the endogenous aspects of technological progress. In the following years, the interest in the neo-classical theory waned, at least in neoclassical circles and this led Fisher to observe in a long and detailed review on developments in macroeconomics in the 1970s and '80s for the *Economic Journal* that 'after a rapid development in the '50s and '60s, the theories of economic growth and capital have received little attention for almost two decades' (Fisher, 1988, p. 37).

There has been a soaring revival of interest only recently. Endogenous growth theory is spreading and many mechanisms have been adopted by neo-classical authors to make growth an endogenous process. As it is not our aim to supply a review nor a classification but rather to analyse the evolution of the idea of endogenous growth within the neo-classical research programme, we shall limit ourselves to a consideration of the essential elements in Lucas' and Romer's approach, as their contributions remain theoretical landmarks (Jones and Manuelli, 1997).

Stiglitz (1988) observed that all progress in growth theory must overcome two types of obstacle. The first has to do with the mathematical constraints which characterise every dynamic model, constraints which generally are not to be found in static analysis. The second is of a different kind and calls into question the view of economic processes underlying the analytical structure. Each advance in economic theory requires that new intuitions be elaborated which give an economic significance to advances on the analytical level. We shall apply this interpretative key to the new growth theory as elaborated by Romer and Lucas.

As regards the analytical aspect, it is Lucas (1997) who clearly defines the essential ingredient in the new approach to economic growth models. To obtain an endogenous growth model, in the present sense of the expression, the marginal product of the accumulated factor must be higher than the interest rate. If this is not the case, the sector which produces the accumulable factor is not sufficiently productive to guarantee long-run product growth. As Lucas observes: ‘What lesson can we draw from the failure of the neo-classical model? I think there are two. First, the villain is the Law of Diminishing Returns. It is this feature that makes it hard to get sustained growth in a model of a single economy … We have to find a way to repeal this law, theoretically’ (Lucas, 1997, p. 68).

In other words, returns on investment should not decline with accumulation but be independent of it and also sufficiently remunerative. The crux of the problem, in an intertemporal context in which the usual representative agent maximises his utility, is that the conditions on technology, and thus on the supply side, must be defined appropriately so that the return of the accumulated factor is not zero with the growth process. What Lucas calls for, and what the new models will supply in many ways, is some modification of the traditional aggregate function of production which moves in the direction required, which is that of contrasting the effect of decreasing returns. So we must turn to the hypothesis of increasing, or at least, non-decreasing, returns to scale, re-elaborated as necessary to satisfy the criterion of balanced growth.

If this is the common analytical skeleton of the new class of growth models, there are a number of ways of reaching the required result according to the specifications of the way in which knowledge is used, produced and distributed in the economy. Romer and Lucas, and later many others, could not but return to reflect on the issue, but barely-sketched during the 1960s, concerning the concept of knowledge as an autonomous factor of production which has particular characteristics that are totally different from other traditional factors of production. To quote Lucas again and his criticism of the Solow model:

> The neo-classical model focuses on the capital accumulation decision, but it is growth in ideas – not merely capital – that drives the system. This observation suggests a shift of focus from decisions on capital accumulation to decisions that determine rate of production of ideas (Lucas, 1997, p. 68).

With economic growth now depending on the accumulation of ideas and no longer on the accumulation of physical capital, the points in common between the old and new theories of endogenous growth are plain. The two approaches share the same plan even though the end results were to be very different. What is clear in particular is that, albeit in a different form, both Romer and Lucas pick up again the concept of increasing returns to scale, or rather, they give a different reply from Solow’s to the problem of inserting increasing returns into a general equilibrium model.

Without going into the analytical details of the models, the basic idea behind Romer’s approach is outlined clearly and concisely in a short essay with the pertinent title: *Are Non-Convexities Important to Understanding Growth?* (1990). His thesis can be summarised in the following terms: if technological progress consists in the accumulation of new ideas and immaterial goods, then it is inevitable that we should turn to the principle of increasing returns to scale. Knowledge in fact is a non-rival factor which has a high production cost but it can be replicated without excessive cost. Here we find Shell’s, and before him, Kaldor’s, idea that technology is a public good and that consequently the function of production is characterised by increasing returns to scale. In Romer’s words:

> The oldest question in Economics is what causes growth. One of the oldest conjectures, built into Adam Smith’s story of the pin factory, is that non-convexities are important for growth […]. We now know how to fit this kind of effect into an aggregate growth model, and we can already see that these models generate many theoretical possibilities (Romer, 1990, p. 98).

But as we have already seen in Shell, technology as a public good cannot grant endogenous growth. Romer’s step forward consists in assuming that increasing returns are due to an externality effect linked to the accumulation of capital on the part of the single firm, that is, that it has the characteristics of a public good but is produced privately. As new discoveries are non-excludable, the single firm produces knowledge from which all the other firms can also benefit thanks to the circulation of information. This externality effect, practically irrelevant at a micro-economic level, becomes the decisive factor on the macro-economic level and thus also for growth theory. To ensure that the long-run growth rate is constant it is necessary to introduce an extremely particular condition. The aggregate effect due to capital accumulation must exactly balance the tendency to decreasing returns which are found at micro level. Otherwise, it is impossible to achieve a balanced growth situation and the economic system tends to explode. Here we find once again, albeit in a different form, the problem of instability which seems to characterise dynamic linear models from Harrod on.

Lucas' model is less innovative than Romer's on the interpretative level but nonetheless has its roots in earlier endogenous growth literature. Lucas recognises his debt to Uzawa and presents his 1988 model as a variant of Uzawa's (1965). In his words: 'In 1965, Uzawa showed that a growth model based on human capital accumulation, without diminishing returns, can produce sustained growth without the *deus ex machina* of exogenous technical change' (Lucas, 1997, p. 68), Lucas gives Uzawa's model a microeconomic syntax which was totally missing from the original model. The research sector is replaced by the notion of human capital, but, above all, the accumulation of human capital depends linearly on the time which each worker dedicates to study and training. With these modifications of interpretation but with the same analytical structure, long-run growth rate per man becomes endogenous in the sense that it depends on the fundamental parameters of the economic system, such as preferences and production function parameters.

To sum up, both Romer and Lucas completely re-orientate the neo-classical growth theory, shifting the focus of the analysis from material resources to immaterial, picking up again the thread of the endogenous growth theory of the 1960s. Once on this road, the new growth theory also found itself with the problem of how to incorporate increasing returns into the theory of general economic equilibrium. Their reply was to completely set aside the problem of distribution, that is, the basic idea of the neo-classical school that there is a correspondence between the quantity of a factor and its price, seeing as this was the obstacle which had to be overcome.

3.7. CHANGES TO THE RESEARCH PROGRAMME

The new growth theory has had remarkable success in giving new energy to the neo-classical research programme, going back to the fundamental question of the factors which determine economic growth and abandoning the static vision of competitive economic equilibrium. The blossoming of the new models was made possible by a change which concerned the research programme's hard core. The principle of decreasing returns proved to be a barrier to the understanding of growth. It was substituted by a new proposition which made it possible to view long-run growth as an endogenous fact, that is, tied to the behaviour of economic agents. The core of the neo-classical research programme on growth now included the following propositions:

- HC1) Growth theory concerns itself with the conditions under which an economy grows in steady-state conditions.
- HC2) The dynamics of an economic system is determined by the accumulation of the factors of production.
- HC3) The supply side of the economy is described by an aggregate production function which allows complete and immediate substitutability of the factors of production which are typically labour and capital.
- HC4) The aggregate production function is characterised by constant returns to scale on the accumulated factor.
- HC5) Growth is determined by the accumulation of immaterial capital.

That there should have been changes to the hard core of the research programme begs the question whether there has been a progressive change in the research programme and therefore an internal adjustment, or whether we are faced with a new programme which relegates the Solovian model to the attic.

When Lakatos describes the creative shifts of a research programme it is clear he is referring to changes in the positive heuristic while the hard core remains intact. This could suggest that with the new endogenous growth theory we do not have a shift forward of the programme as in the case, for example, of the endogenisation of savings in the 1960s or of the vintage models, but a whole new growth research programme. It could be argued that the new growth theories have in fact created an alternative programme to the dominant Solovian program.

But this rigid application of Lakatos' approach would, however, lead us in the wrong direction. We could hardly say that economists involved in the endogenous growth research project are outside the neo-classical research programme, in the usual meaning of the term. The endogenous growth approach shows a large degree of continuity in the neo-classical research programme and the greatest effort on the part of Lucas and his school is to prove the superiority of the neo-classical research programme over rival programmes (Lucas, 1988).

This seeming paradox springs from the difficulty of precisely identifying the elements of a research programme and can easily be overcome if we use Remenyi's suggestion, when discussing the application of Lakatos' methodology to economics (Remenyi, 1979), that the categories of a research programme be made more flexible by introducing the idea of demi-cores. For Remenyi, each research programme generates in its development a series of specialties and sub-disciplines that have common features, each of which is characterised by its own core, named demi-core. The protective belts of sub-programs can overlap and, although their demi-cores may be distinct, they share common elements mediated through the hard core. For Remenyi 'the demi-core is to the sub-discipline what the hard core is to the MSRP' (Ibidem, p. 33). The important point is that the dynamics of a research programme is determined by the evolution of the sub-programmes it can generate which map out the heuristic path of the programme. He states: 'It is a fundamental result of the theory of core demi-core interaction that the number of demi-core is not constant over time' (Ibidem, p. 34), but the heuristic of the core continually generates specialties and demi-core which testify to the vitality of a research programme.

This elaboration of Remenyi's allows us to get out of the impasse into which a rigid application of Lakatos' methodology led us. The new growth theory may be considered not so much a new research programme but rather a new articulation of the neo-classical programme on economic growth which has led to the formation of a new demi-core capable of filling the gaps in the previous one though belonging to the same research programme.

Which element distinguishes the new demi-core from the previous one? For the new growth theorists the answer is plain: the fundamental limit of the Solovian approach is to be found in the fact that it was lacking a micro-economic theory of technological progress. The new theory of endogenous growth may therefore be considered part of a more complex and ambitious project carried on by the new neo-classical macroeconomics to rethink macroeconomic analysis on the basis of the fundamental assumption that individual agents make optimal choices in markets which are linked to each other and that these markets reach some sort of equilibrium. Equipped with tools from general equilibrium economics theory, growth economists tried to solve the old problem of giving a serious microeconomic foundation to macroeconomics. In the 1970s, the economists of the new classical macroeconomics school developed a theory of the economic cycle as optimal deviation of output around a trend; in the following decade they attempted to explain the trend itself. As growth theory economists have often pointed out, their contribution was that they successfully inserted the old idea of increasing returns into a general economic equilibrium context.

3.8. FURTHER ASPECTS OF LINEAR GROWTH MODELS

In the previous section we saw that in some of the recent NGMs the dominant idea is to drop the non-produced factors in the production function so as to avoid any sources of decreasing returns assuming linearity in the production function. Romer obtains this result assuming that the determinants of technological progress are non-rival factors and Lucas by

assuming that each producer benefits from the level of human capital. Besides the methodological implications that we analysed in the preceding section, the linearity in the production function produces some crucial consequences that are worth considering.

The first is the troublesome result that a non-accumulable factor like labour does not have any role in production, for otherwise the growth rate would depend upon the population level (scale effect), as with Arrow's model. This is quite evident in the simple version of this kind of model, the so called *AK* model (Rebelo, 1991), which starts from the premiss that there is a single factor of production, capital, and all the other factors are simply eliminated by hypothesis. In turn, Romer and Lucas face this crucial problem only for very specific assumptions, arguing that the technological parameter depends on the economy's average capital (human or physical) per worker, rather than aggregate capital stock, *H* or *K*. The upshot of this solution is the totally unrealistic assumption that the production function no longer depends on *L*, the labour force employed.

If it is hard to think of an economic process without non-accumulable factors such as labour or natural resources, the difficulties for growth theorists are even tougher. One striking implication of linearity assumptions is that labour's share of income becomes asymptotically zero, which reopens the issue of what factors govern income distribution (Bertola, 2000). It seems that the price to be paid in order to make the rate of growth endogenous is to abdicate to the traditional theory of income distribution in which relative factor prices reflect relative scarcity and the amount which each factor obtains from the national product is determined by technology and relative factor endowments.

The second interesting feature of NGMs is tied to the previous one in the sense that, if income distribution no longer depends upon the endowment factor, then it is possible to consider profit as the relevant exogenous variable that makes the micro foundations of the growth process possible. If this is true, recent development can be characterized by a partial return to ideas that were prominent in growth theory prior to Solow's model (Kurz and Salvadori, 1998b, 1999). These models adopt a simplified version of the classical notion of production as a circular flow and of profit as a surplus product. As in the classical tradition there is no limit to growth because it is generated within the growth process and the growth rate is endogenously determined assuming a relationship between the rate of growth and the rate of profit. In the simple *AK* model, since the consumption good is produced only by means of capital and the saving rate is constant, the growth rate is simply the profit rate multiplied by the saving rate (Kurz and Salvadori, 1998b, p. 76).

The genuinely novel element consists in the fact that profit is determined by technology alone and the growth rate of the system is then determined by the saving–investment equation: the larger the propensity to accumulate, the larger the growth rate. In the former approach, the endogenous aspect of economic growth referred to various institutional, social and economic mechanisms that were able to induce economic change. In NGMs these mechanisms are generally based on some special technological relationship placed in the equation that describes the accumulation of the relevant factor in the long run.

3.9. CONCLUDING REMARKS

In this paper I have viewed the development of neoclassical growth theory from a Lakatosian perspective in order to evaluate the relevance of Lakatosian ideas for economics. The main conclusion that emerges is that the rational reconstruction of neo-classical growth theory fits the facts.

In the first place, it has been possible to characterize neo-classical growth theory as a genuine research programme based on capital accumulation and exogenous progress. The evolution of this programme could be described in terms of Lakatosian problem shifts that have concerned essentially the question of making saving and technological progress endogenous variables.

Based on the key hypothesis of exogenous technical progress, the neo-classical programme first flourished and later, in the 1970s, stagnated. The current rebirth was determined by the return to earlier frameworks that were primarily based on increasing returns to scale bound up with the process of knowledge accumulation developed in the sixties. In Lakatosian terms, the new models could be interpreted as a change that went beyond the impact on the positive heuristic of the Solovian approach in that they effected a change in the core propositions of the dominant programme. Thus, using Remenyi's suggestion, they resulted in the development of a new demi-core in the neoclassical tradition that in certain ways mark a return to Harrod and to traditions that go further back in the history of economic thought.

NOTES

1. See Backhouse (1998) for a discussion of the Lakatosian methodology applied to economics.
2. Following Backhouse we consider only the propositions which are relevant to growth theory. For a complete presentation see Remenyi (1979) and Weintraub (1985)

3. The idea that the economic system is driven by exogenous factors precedes Solow's contribution and it appears in several authors, see for instance Marshall and Cassel (Kurz and Salvadori, 1999, pp. 246–47)
4. For a more in-depth analysis of this point see Pomini (2000).

4. Endogenous growth in a multi-sector economy

Giuseppe Freni, Fausto Gozzi and Neri Salvadori

4.1. INTRODUCTION

There are at least three different approaches to endogenous growth (see Jones and Manuelli, 1997). Two include non-convexities or externalities or both. The third relies on convex models of growth in which, properly interpreted, the two welfare theorems hold (e.g., Jones and Manuelli, 1990 and Rebelo, 1991). The models in this last strand of literature are characterized by the fact that production is not limited by primary resources and hence the equilibrium paths can show endogenous growth. For the sake of simplicity we call 'convex models' the models whose equilibria satisfy the conditions of the two Fundamental Theorems of Welfare Economics, and subdivide them into 'bounded models' (those whose feasible paths are limited by the availability of natural resources) and 'unbounded models'. In the last fifteen years, models with explicit consumption and a production side in which 'goods are made out of goods alone'[1] have been widely used in the new growth theory, especially in that approach to endogenous growth based on the assumption that all production factors are reproducible (Lucas, 1988; Rebelo, 1991). It could be argued that any mechanism found in the literature to make sustained growth possible has essentially involved the assumption that there is a 'core' of capital goods whose production does not require (either directly or indirectly) non-producible factors. In fact the reduced form of most endogenous growth models is linear or asymptotically linear in the reproducible factors (see for example, Frankel, 1962; Romer, 1986, 1987, 1990).[2] Consequently, in the endogenous growth literature, static analysis is mainly centered around the concept of the 'balanced growth path', which, in this context, performs the role played by the stationary state in convex bounded models.

It is common in the literature on growth to study one- or two-sector models. The exceptions are some convex models. Recent contributions to *n*-

sector unbounded growth models have been provided by Dasgupta and Mitra (1988), Dolmas (1996), Kaganovich (1998), Ossella (1999), and Freni, Gozzi and Salvadori (2001, hereafter FGS). In FGS we studied a multi-sector 'AK model' in continuous time. We provided an existence result for optimal strategies, a set of duality results, and a complete classification of the price-supported steady states of the model. In doing so we used a number of assumptions (for the full list of assumptions used in FGS see Section 4.3 below). A consequence of combining these assumptions with the lack of primary resources that characterizes the 'AK' model was that for positive rates of growth (low discount rates) the structure of prices in the unique steady state of the model turned out to be invariant with respect to preferences. Hence the independence of the interest rate from preferences that holds in the one-sector 'AK' model carries over to the multi-sector model we studied in FGS. But this result is not confirmed when negative growth rates (i.e. high discount rates) are considered. On the contrary, we envisaged two other regimes, in one of which the prices in the steady state depend dramatically on the discount rate. This proves that the introduction of a multiplicity of sectors entails problems which are not visible in a one-sector model even if stability is not taken into account. However, the multi-sector linear growth model in FGS still has a very simple static structure and predicts that the long-run rate of growth of an economy does not depend on the initial conditions. Thus one could wonder if path dependence can be obtained in a linear endogenous growth framework or whether increasing returns and/or imperfectly competitive markets are required to obtain it. In the present paper we will discuss removing some of the assumptions made in FGS in order to show that the structure of the steady-state set of the 'AK' model can be considerably enriched. Moreover, we provide several interpretative elements, relevant in dealing with convex unbounded models.

The various building blocks of the model presented in FGS have different origins. In particular, the production side of the model has a clear classical flavour (see Kurz and Salvadori, 1995) and is very close to the production side of the von Neumann model, in which commodities are produced out of each other because '[w]age costs are not considered as such, for laborers are not separately considered any more than are farm animals' (Champernowne, 1945, p. 12). This structure hints at a 'technological' theory of the long-run rate of interest, that 'appears as the natural and optimum rate of organic expansion of the system, and depends on the technical processes of production which are available' (Ibidem). Nevertheless, the model has also a Ramsey-like preference side in which the optimal behaviour of the representative agent determines the system's saving rate. Since we maintain the hypothesis that the behaviour of the representative agent does not affect the technical conditions of production, our rational agent cannot be a worker, because the real wage is still ‘whatever is needed to persuade people to work’ (Ibidem, p. 16). Therefore, it is ‘[t]he question of consumption by the propertied class’ (Ibidem) that properly arises in the model with an intertemporal utility functional. However, even under this ‘representative capitalist’ interpretation, the introduction of explicit consumption creates a tension in the model as regards the forces determining the long-run rate of interest.[3] Both considerations of technology-and-cost arbitrage and preferences concur in determining the growth rate, the profit rate and the relative prices that prevail in the long-run equilibrium. Nevertheless, in the model there is some space for the ‘classical’ opinion that ‘even if part of the income from property were spent on consumption, and not saved, the rate of interest would not necessarily be much affected: it might still be *approximately* equal to the greatest expansion rate that *would* have been possible *if* all income from property had been saved’ (Ibidem).

Although the production framework used in FGS is the ‘simple linear production model’ (Gale, 1960, p. 294) which excludes both joint production and choice of technique, the complete set of results we summarize in the Classification Theorem provided in that paper is novel. The reason is that the study of multi-sector closed linear models of production with explicit consumption has been confined almost exclusively to the discrete-time framework (see McFadden, 1967; Atsumi, 1969 and the more recent literature mentioned above), while we study a continuous-time model. As a consequence, some kinds of complications which in discrete time can be avoided at first (see however Atsumi, 1969, p. 270), arise from the beginning. These complications are generally connected with joint production and decomposability, but in a continuous-time model they can also be connected with the way fixed capital is formalized in order to avoid an infinite number of commodities (depreciation by evaporation).

The paper is organized as follows. In Section 4.2 we present the model in a general format whereas in Section 4.3 we present the results obtained by FGS in a restrictive setting. Section 4.4 is devoted to a comparison with the von Neumann–Sraffa–Morishima models developed mainly in the 1960s and 1970s. Section 4.5 clarifies the extensions of the results by FGS which, on the basis of the comparison, are expected to be easily obtained and what should require more effort. It is also devoted to clarifying the differences in the analysis connected with relaxing some assumptions, allowing decomposability of the input matrix, joint production, and multiple consumption goods. Section 4.6 provides some conclusions.

4.2. A CLASS OF CONVEX MULTI-SECTOR ENDOGENOUS GROWTH MODELS IN CONTINUOUS TIME

The commodity space is finite and the technology is stationary. Production consists in combining the productive services from the stocks to generate flows that add to the existing stocks. Decay and consumption, on the other hand, drain away the stocks. The production set is generated by a finite number *m*, $m \geq 1$, of independent activities (or processes), each of which can be run at any scale of operation. Hence, process *j* (j = 1, 2, ..., *m*) can be represented by a pair of *n*- dimensional input–output vectors:

$$\mathbf{a}_j^T \rightarrow \mathbf{b}_j^T ,$$

where $\mathbf{a}_j^T\mathbf{e}_i \geq 0$ is the amount of productive services from stock *i* that process *j* uses at the unitary level of activation and $\mathbf{b}_j^T\mathbf{e}_i \geq 0$ is the flow of commodity *i* produced by the same process at the same level of activation. Thus, we can summarize the production processes with a pair of *m*x*n* non-negative matrices:[4]

$$\mathbf{A} \rightarrow \mathbf{B} .$$

For the sake of simplicity, we assume that the rate of decay of the stocks in production is given by a single constant δ_x, $\delta_x \geq 0$. Stocks not used in production are ‘stored’. Stored commodities decay at the rate δ_z, $\delta_z \geq 0$. To avoid jumps in the stocks we conceptualize ‘disposal’ as a storage process with $\delta_z \geq \delta_x$. Each stock can be produced and no productive service from primary factors is relevant in the system. The way we handle storage (or disposal) processes implies that we assume that, in addition to the *m* processes formally included in the technology, free disposal activities for the productive services of the stocks are available.

We give a basic characterization of the technology by means of the following two classical von Neumann-like assumptions (Gale, 1960, p. 311):

[HP1] Each column of matrix B is semipositive

This assumption means that all commodities are reproducible and therefore there is no primary factor.

[HP2] Each row of matrix **A** is semipositive

This assumption means that no process can be activated without using (the service of) some commodity as an input. It therefore implies that for each $t \geq 0$ the intensity levels of the production processes are bounded from above by the existing stocks. It is convenient to assume that the system can grow at a positive rate since this case is the most interesting one from an economic point of view:

[HP3] $\exists\, \mathbf{x} \geq \mathbf{0}, g > 0 : \mathbf{x}^T \left[\mathbf{B} - (\delta_x + g)\mathbf{A}\right] \geqq \mathbf{d}^T$, where $\mathbf{d}$ is any non-negative vector proportional to the vector of consumed commodities.

FGS and Gozzi and Freni (2001) concentrate on the case in which there is no joint production in flow outputs. That is:

[FGS1] Each row of matrix **B** has one and only one positive element

Note that in this case each positive element of **B** can be normalized to 1 without loss of generality. Under Assumption [FGS1] processes can be unambiguously linked to industries. If, moreover, each industry has one production process, then the input/output matrices are square and we will have:

[FGS2] $m = n$ (i. e. **A** is square and **B** = **I**)

Another assumption about production we sometime use is:

[FGS3] $\mathbf{x}^T \left[\mathbf{B} - (\delta_x + g)\mathbf{A}\right] \geqq \mathbf{0}^T, g > -\delta_x, \mathbf{x} \geq \mathbf{0}, \mathbf{x} \neq \mathbf{0} \Rightarrow \mathbf{x}^T \mathbf{B} > \mathbf{0}^T$

It means that the services from each stock enter directly or indirectly into the production of each commodity (see Kurz and Salvadori, 1998c, pp. 95–7). In particular, if Assumptions [FGS1] and [FGS2] hold, [FGS3] is equivalent to the assumption:

[FGS4] The square matrix **A** is indecomposable

As mentioned above, the way disposal is conceptualized implies the assumption

[HP4] $\delta_z \geq \delta_x$

The discount rate, $\rho \in \Re$, and the instantaneous utility function, *u*, describe the preference side of the economy because, as is quite common in the new growth theory, we are dealing with a single-consumer economy whose preferences can be represented by a utility functional $U(\mathbf{c}(t))$ with the form

$$U(\mathbf{c}(t)) = \int_0^\infty e^{-\rho t} u(\mathbf{c}(t))dt\ ,$$

where

$$u(\cdot) : \Re_+^{c_m} \to \Re \cup \{-\infty\}, 1 \leq c_m \leq n\ .$$

Note that if $c_m < n$, then there are $n - c_m$ pure capital goods. Since the production side of the model is linear, the whole model is homogeneous if the preference side is so. We therefore assume that the utility function is the usual iso-elastic one:

[HC1] $u(\cdot) = \dfrac{1}{1-\sigma}[v(\cdot)]^{1-\sigma}$ for $\sigma > 0,\ \sigma \neq 1$, $u() = \log[v(\cdot)]$ for $\sigma = 1$;

[HC2] $v(\cdot) : \Re_+^{c_m} \to \Re_+$ is increasing, concave and homogeneous of degree one. If $c_m > 1$, then $v(\cdot)$ is strictly concave.[5]

Preferences are fully described by ρ, σ, and function $v(\cdot)$. In the following we will say that $v(\cdot)$ describes the preferences concerning consumption at a given moment in time, whereas ρ and σ describe preferences concerning distribution of consumption (and saving) over time. Note that if $c_m = 1$, the utility function $v(c) := c$ is equivalent to any other. Hence, if [6]

[FGS5] $c_m = 1$,

then parameters ρ and σ completely describe the preference side of the model.

Let **s** be the *n*x1 vector of stocks, **x** be the *m*x1 intensity vector and $\hat{\mathbf{c}}$ be the *n* x 1 vector obtained from the consumption vector **c** by adding a zero component for each pure capital good. The evolution of the stocks is given by the following differential equation:

$$\dot{\mathbf{s}}^T = \mathbf{x}^T\mathbf{B} - \delta_x \mathbf{x}^T\mathbf{A} - \delta_z(\mathbf{s}^T - \mathbf{x}^T\mathbf{A}) - \hat{\mathbf{c}}^T$$

with the constraints

$$\mathbf{x}^T\mathbf{A} \leq \mathbf{s}^T, \mathbf{x} \geq \mathbf{0}, \mathbf{c} \geq \mathbf{0}.$$

The common approach to the analysis of competitive equilibria in the above setting is through the extension of the first and second welfare theorems for finite dimensional economies. This strategy leads to investigate the link between the competitive equilibria of the system and the solutions, if there are any, to the problem:

$$\begin{aligned} V(\mathbf{s}^*) &= \sup \int_0^\infty e^{-\rho t} u(\mathbf{c}(t))dt \\ \dot{\mathbf{s}}^T &= \mathbf{x}^T\mathbf{B} - \delta_x \mathbf{x}^T\mathbf{A} - \delta_z(\mathbf{s}^T - \mathbf{x}^T\mathbf{A}) - \hat{\mathbf{c}}^T \\ \mathbf{x}^T\mathbf{A} &\leq \mathbf{s}^T,\ \mathbf{x} \geq \mathbf{0},\ \mathbf{c} \geq \mathbf{0},\ \mathbf{s}(0) = \mathbf{s}^* \geq \mathbf{0}\ \textit{given}. \end{aligned} \qquad \text{(P)}$$

We will therefore be interested in the existence and characterization of the paths for which the problem (P) has a solution. Moreover, since under the above assumptions (P) is a homogeneous program, our interest will lie also in the existence and characterization of the special paths that solve (P) and enjoy a steady state structure. These paths provide the simplest reference point for the analysis of the asymptotic behaviour of non-stationary paths.

In studying the optimal control problem (P), Hamiltonian formalism is often used to introduce the 'price' variables. From an economic point of view, this is a particularly significant procedure because it leads to the introduction of competitive prices. Indeed, what are defined as competitive paths are simply stock-price paths supporting the maximized Hamiltonian (see Cass and Shell, 1976). In our context, since the optimal control problem (P) is autonomous, the discounted Hamiltonian is used. It is given by:

$$H^D(\mathbf{s},\mathbf{v}) = \max_{\mathbf{c}\geq\mathbf{0}}\left[u(\mathbf{c}) - \hat{\mathbf{c}}^T\mathbf{v}\right] - \delta_z \mathbf{s}^T\mathbf{v} + \max_{\substack{\mathbf{x}^T\mathbf{A}\leq\mathbf{s}^T \\ \mathbf{x}\geq\mathbf{0}}} \mathbf{x}^T[\mathbf{B} - (\delta_x - \delta_z)\mathbf{A}]\mathbf{v}\,.$$

The linear programming problem involved in the definition of $H^D(\mathbf{s},\mathbf{v})$ requires the existence of a vector $\mathbf{q}^*(\mathbf{s},\mathbf{v}) \in \Re^n_+$ which is a solution to the dual problem and can be interpreted as the vector of the equilibrium 'rental' rates for the use of the stocks. Hence, if $\mathbf{c}^*(\mathbf{v}), \mathbf{x}^*(\mathbf{s},\mathbf{v})$ indicate a set of controls solving the max problems involved in the definition of $H^D(\mathbf{s},\mathbf{v})$, then the set of paths $[\mathbf{s}(t), \mathbf{v}(t), \mathbf{c}^*(\mathbf{v}(t)), \mathbf{x}^*(\mathbf{s}(t),\mathbf{v}(t)), \mathbf{q}^*(\mathbf{s}(t),\mathbf{v}(t))]$ì, which satisfy appropriate continuity properties and solve

$$\dot{\mathbf{s}} \in \partial H^C_{\mathbf{v}}(\cdot,\cdot) \tag{1}$$

$$\dot{\mathbf{v}} - \rho\mathbf{v} \in -\partial H^C_{\mathbf{s}}(\cdot,\cdot) \tag{2}$$

is called a *competitive program*, i.e. a critical point for the problem (P). A transversality condition is then involved in the extension of the first and second welfare theorems to infinite horizon economies. The first welfare theorem, in particular, will state that absolutely continuous competitive paths with nonnegative prices that satisfy a suitable transversality condition are optimal. It is well known, however, that a full converse of this result does not hold due to the possibility that absolutely continuity of prices cannot be granted for stock paths hitting the nonnegativity boundary. Moreover, prices supporting non-interior stock paths can fail to exist altogether (examples of this phenomenon are provided by FGS, Appendix D).

In closing this section, we should point out that there are some convex endogenous growth models that cannot be reduced at once to the present framework, a few of which are mentioned here. First, models with pure consumption goods and/or non-reproducible resources that are not essential

(see e.g. Bose, 1968; Weitzman, 1971; Jones and Manuelli, 1990; Rebelo, 1991). Second, models with adjustment costs (see e.g. Dolmas, 1996; Ladron-de-Guevara, Ortigueira and Santos, 1999). Third, models in which the technology is not polyhedral (e.g. Kaganovich, 1998, Jensen, 2000). Finally, models with an infinite dimensional commodity space (Boldrin and Levine, 2002).

4.3. STEADY STATES IN A 'SIMPLE' MULTI-SECTOR AK MODEL

In FGS Assumptions [HP1]–[HP3], [FGS1]–[FGS5], [HC1]–[HC2] hold. Moreover, it is also assumed that all commodities are available at time 0 and that the (unique) consumption good enters directly in its own production. That is, if the consumption good is commodity 1,

[FGS6] $\mathbf{s}^* > \mathbf{0}$,

[FGS7] $a_{11} > 0$.

The first of these two assumptions implies that there is an admissible solution to problem (P) with a positive $\mathbf{s}$ for each $t > 0$. The second assumption guarantees that $\mathbf{s} > \mathbf{0}$ for each $t > 0$ in each optimal solution starting at any $\mathbf{s}^* > \mathbf{0}$. This result is relevant since if $\mathbf{s} > \mathbf{0}$, then $\mathbf{q}$ is bounded and, therefore, prices $\mathbf{v}$ cannot jump. FGS first prove that an optimal path exists if and only if

[HE] $\sigma\Gamma > \Gamma - \rho$

where

$$\Gamma = \sup\left\{g : \exists\, \mathbf{x} \geq \mathbf{0} : \mathbf{x}^T\left[\mathbf{B} - \left(\delta_x + g\right)\mathbf{A}\right] \geq \mathbf{d}^T\right\}$$

and $\mathbf{e}_1$ is the first unit-vector: a vector proportional to the vector of consumed commodities.[7] Under the assumptions maintained by FGS it turns out that $\Gamma = \lambda_{PF}^{-1} - \delta_x$, where λ_{PF} is the Perron–Frobenius eigenvalue of matrix $\mathbf{A}$.

The above existence theorem is completed with two theorems concerning the optimality conditions for the problem at hand, that are the extensions to the present framework of standard results holding in smooth-bounded models. In particular it is proved (*i*) that a competitive program is optimal if the usual condition that the value of the stocks converges to zero holds and (*ii*) that for each optimal solution to problem (P) it is possible to find prices and rentals paths such that the optimal solution is also a competitive program whose value of the stocks converges to zero. For formal statements and corresponding proofs see the original paper by FGS.

As a step towards the study of the dynamics of the system, FGS analyzed the *steady-state optimal solutions* to problem (P), which are defined as optimal solutions $[\mathbf{s}(t), \hat{\mathbf{c}}(t)), \hat{\mathbf{x}}(t)]$ to problem (P) for which there is a real number g, a real number c_0, and a non-negative vector $\mathbf{x}_0$ such that $\hat{\mathbf{c}}(t) = c_0 e^{gt}\mathbf{e}_1$, $\hat{\mathbf{x}}(t) = \mathbf{x}_0 e^{gt}$, where $\mathbf{e}_1$ is the first unit vector. The definition of steady-state optimal solutions does not state that the supporting relative prices are constant over time.[8] However, FGS have proved that for each steady-state solution there are a price path and a rental path such that

$$\mathbf{v}(t) = \mathbf{v}(0)e^{-g\sigma t}, \; \mathbf{q}(t) = (\rho + \delta_z + g\sigma)\mathbf{v}(0)e^{-g\sigma t}.$$

We will refer to these prices and rentals as *steady-state price-rental paths*. On the basis of these and other results, FGS reduce the problem of the optimal steady states to the analysis of finding scalars $g \in [\frac{1}{\sigma}(\lambda_{PF}^{-1} - \delta_x - \rho), \lambda_{PF}^{-1} - \delta_x)$ and c_0 and vectors $\mathbf{s}_0$, $\mathbf{x}_0$, and $\mathbf{v}_0$ such that:

$$\mathbf{e}_1^T \mathbf{v}_0 = c_0^{-\sigma} \tag{3a}$$

$$[\mathbf{I} - (\rho + \delta_x + g\sigma)\mathbf{A}]\mathbf{v}_0 \le \mathbf{0} \tag{3b}$$

$$\mathbf{x}_0^T [\mathbf{I} - (\rho + \delta_x + g\sigma)\mathbf{A}]\mathbf{v}_0 = \mathbf{0} \tag{3c}$$

$$\mathbf{x}_0^T \mathbf{A} - \mathbf{s}_0^T \le \mathbf{0} \tag{3d}$$

$$\mathbf{x}_0^T [\mathbf{I} - (g + \delta_x)\mathbf{A}]\mathbf{v}_0 = c_0 \mathbf{e}_1^T \mathbf{v}_0 \tag{3e}$$

$$\mathbf{x}_0^T [\mathbf{I} - (g + \delta_x)\mathbf{A}] \ge c_0 \mathbf{e}_1^T \qquad \text{for } g + \delta_z \ge 0 \tag{3f}$$

$$\mathbf{s}_0^T - \frac{1}{g + \delta_z}\{\mathbf{x}_0^T [\mathbf{I} - (\delta_x - \delta_z)\mathbf{A}] - c_0 \mathbf{e}_1^T\} \ge \mathbf{0} \quad \text{for } g + \delta_z < 0 \tag{3g}$$

$$c_0 > \mathbf{0}, \; \mathbf{v}_0 \ge \mathbf{0}, \; \mathbf{x}_0 \ge \mathbf{0}. \tag{3h}$$

Optimal steady state solutions with steady state support are relevant also since in these states some relevant concepts, such as that of ‘real rate of profit’ or ‘growth rate’, can be defined. The literature on growth often refers to steady states in order to convey some macroeconomic insights. Lucas (1988. p. 11), for instance, refers freely to steady state concepts as ‘the rate of growth’ or the ‘real rate of profit’ under the explicit assumption of a fast convergence to the steady state.

As a matter of fact there are profitability concepts which can be used with reference to optimal solutions even if they are not steady states. It is always

possible to deal with the ‘own rate of return of commodity *i*’ as the rate of profit which can be obtained by an investment measured in commodity *i* getting a revenue measured in commodity *i* (see, for instance, Malinvaud, 1953). FGS show that in any optimal solution[9]

$$r_i(t) = \rho - \frac{\mathbf{e}_j^T \dot{\mathbf{v}}(t)}{\mathbf{e}_j^T \mathbf{v}(t)}$$

where $r_i(t)$ is the own rate of return of commodity i at time t. FGS also proved that in any optimal solution the growth rate of consumption equals the ratio of the difference between the own rate of return of commodity 1 and ρ over σ:

$$\frac{\dot{c}(t)}{c(t)} = \frac{r_1(t) - \rho}{\sigma}. \tag{4}$$

In a steady-state solution supported by steady-state prices all the own rates of return are equal to each other so we can call this common rate the ‘real rate of profit *r*'. Similarly, in a steady-state solution all the intensities of operation of processes as well as consumption grow at the same rate so we can call this common rate the ‘growth rate *g*'. Obviously, in a steady state

$$g = \frac{r - \rho}{\sigma}. \tag{5}$$

From equations and inequalities (3) FGS obtained a Classification Theorem, in which three different regimes are envisaged, depending on the values of the parameters involved. The Classification Theorem states a particular relationship between the growth rate g and the rate of profit r. This $r - g$ relationship is drawn in Figure 4.1. In the first regime r is constant and equals $\lambda_{PF}^{-1} - \delta_x$, whereas g varies in the range $(-\delta_x, \lambda_{PF}^{-1} - \delta_x)$. In the second regime g is constant and equals $-\delta_x$, whereas r varies in the range $[\lambda_{PF}^{-1} - \delta_x, a_{11}^{-1} - \delta_x]$. In the third regime r is constant again and equals $a_{11}^{-1} - \delta_x$, whereas g varies in the range $(-\infty, -\delta_x)$. This relationship is not to be confused with another r–g relationship, that is, equation (5). The former depends on technology and on preferences concerning consumption at a given moment in time and does not depend on preferences concerning distribution of consumption over time.[10] The latter, by contrast, depends only on preferences concerning distribution of consumption over time. In a steady state solution r and g are determined by the intersection between these two $r - g$ relationships.

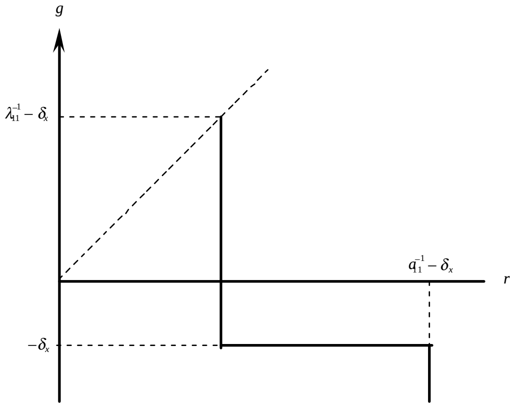


Figure 4.1

The interpretation of the Classification Theorem is simple. If $g > -\delta_x$, then all commodities need to be produced and therefore n processes are to be operated. Thus, the n equations relating prices and rate of profit relative to the operated processes (the no arbitrage conditions) determine both the n – 1 relative prices and the real rate of profit. In this regime prices are proportional to $\mathbf{v}_{PF} > \mathbf{0}$, that is the right eigenvector of matrix **A** corresponding to λ_{PF}. If $g < -\delta_x$, the only process which is relevant is the process producing commodity 1. Since the inputs used by this process are produced jointly by the process itself at a rate larger than the growth rate, all commodities used in the production of commodity 1 except commodity 1 itself have a zero price. In other words, production is reduced to the production of commodity 1 by means of commodity 1 and free goods. Hence, similar to the previous case the equation relating prices and the rate of profit relative to the operated process producing commodity 1 can determine the rate of profit (apart from commodity 1, all commodities which are either produced or stored have a zero price). If $g < -\delta_x$, then once again the only relevant process is that producing commodity 1, and the inputs used by this process are produced jointly by the process itself. Yet this is realized at a rate equal to the growth rate and therefore these commodities (except commodity 1) may have either a positive or a zero price. Those with a positive price cannot be separately produced or stored; their existing stocks can be regarded as stocks of 'renewable' resources for which a growth rate of $-\delta_x$ can be granted in the production of commodity 1.

4.4. A COMPARISON WITH THE VON NEUMANN–SRAFFA–MORISHIMA MODELS

Let us consider the case of $g + \delta_z \geq 0$, taking account of the fact that in a steady state equation (4) holds. Let us substitute r for $\rho + g\sigma$ in inequality (3b) and in equation (3c). Let us drop Assumptions [FGS1]–[FGS4] in order to allow for choice of techniques and joint production and, therefore, substitute matrix **B** for **I** in inequalities (3b), (3g), (3h) and in equations (3c) and (3f). Then we obtain from inequalities (3b), (3g), (3i) and equations (3c) and (3f) exactly the model analyzed in the von Neumann–Sraffa–Morishima literature developed in the 1960s and 1970s, with contributions up to now (Kurz and Salvadori, 1995, summarize the whole approach):

$$[\mathbf{B} - (r + \delta_x)\mathbf{A}]\mathbf{v}_0 \leq \mathbf{0}\,,\ \mathbf{x}_0^T[\mathbf{B} - (r + \delta_x)\mathbf{A}]\mathbf{v}_0 = \mathbf{0}$$

$$\mathbf{x}_0^T[\mathbf{B} - (g + \delta_x)\mathbf{A}] \geq c_0\mathbf{e}_1^T\,, \mathbf{x}_0^T[\mathbf{B} - (g + \delta_x)\mathbf{A}]\mathbf{v}_0 = c_0\mathbf{e}_1^T\mathbf{v}_0$$

$$c_0 > \mathbf{0}\,,\ \mathbf{v}_0 \geq \mathbf{0}\,,\ \mathbf{x}_0 \geq \mathbf{0}\,.$$

The main difference consists in the fact that in the literature in question there was almost always at least one primary factor called 'labour'. In the cases in which no primary factor was taken into consideration, or the wage of labourers was taken to be zero, an extra inequality was mentioned, which in our case it is certainly satisfied since equation (3a) and inequality (3h) hold:

$$\mathbf{x}_0^T\mathbf{B}\mathbf{v}_0 > 0\,.$$

The first problem to be analyzed is the following. The first regime mentioned in the Classification Theorem is mentioned in the von Neumann–Sraffa–Morishima literature. Actually it is 'the' result that would have been predicted for an economy with stationary relative prices and a wage rate equal to zero. But what about the other regimes? The literature in question has rarely considered negative growth rates. Could one expect such a difference?

The problem arises since for $g < -\delta_x$ the distinction between single production and joint production is not relevant. This section is devoted to clarifying this point as a contribution to understanding the problems involved in multi-sector models. A graphic exposition in terms of two goods will be sufficient here. Consider a joint production process

$$\mathbf{a} \rightarrow \mathbf{b}\,,$$

and let us represent it in a commodity space in which vectors **a** and **b** and vector $\mathbf{b}-(g+\delta_x)\mathbf{a}$ appear. Figure 4.2 depicts four alternative vectors $\mathbf{b}-(g+\delta_x)\mathbf{a}$, depending on the size of *g*: $g' < -\delta_x < g'' < g'''$. In Figure 4.3, instead, there are two single production processes, one producing commodity 1 and one producing commodity 2. Also in this case in the figure we have drawn alternative vectors $\mathbf{b}_i-(g+\delta_x)\mathbf{a}_i$. It is immediately recognized that when $g+\delta_x$ is positive, the cone consisting of convex combinations of vectors $\mathbf{b}_i-(g+\delta_x)\mathbf{a}_i$ (i = 1, 2) includes the positive orthant, whereas for negative values of $g+\delta_x$ the same cone is included in the positive orthant. Both processes are needed to produce the consumption vector at the required growth rate only if the consumption vector is *internal* to that cone. Let us

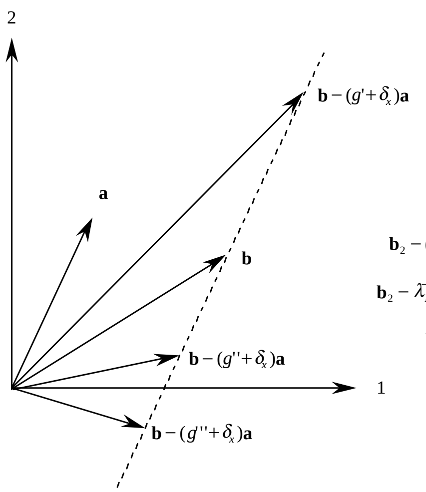


Figure 4.2

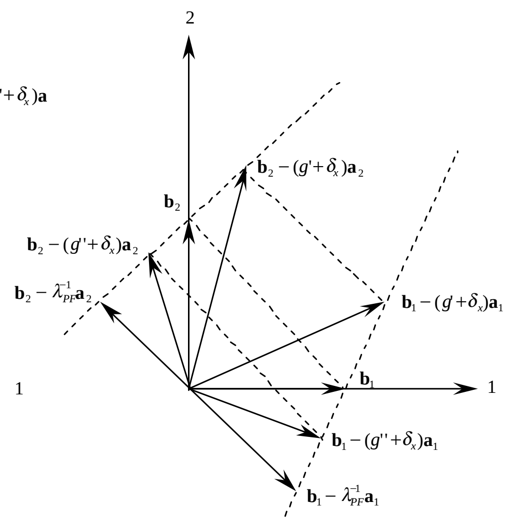


Figure 4.3

focus on the other cases: that in which the consumption vector is *external* to the cone and that in which the consumption vector is *on the boundary* of the cone.

In a two-commodity economy the fact that the consumption vector is external to the above-mentioned cone means that no convex combination of processes can supply an amount of the required consumption without overproduction of a commodity which, as a consequence, has a zero price. Moreover, there is a problem of choice of technique even if the original model contemplated a number of processes equal to the number of commodities: as a matter of fact the model is equivalent to one in which the commodity with a price equal to zero does not exist, whereas the number of processes is unchanged. The same argument is immediately applicable to an *n*-commodity economy: once again some prices need to be zero, a convex

combination of the operated processes meets the components of the consumption vector with a positive price and overproduction of commodities with a zero price.

In a two-commodity economy the fact that the consumption vector is on the boundary of the cone means that there is a convex combination of processes which can supply an amount of the required consumption without the condition that a commodity is overproduced, but this convex combination is actually made up of only one process: no commodity needs to be overproduced, but the no arbitrage conditions determine a number of constraints lower than the number of the prices to be ascertained. Hence we have to study whether the fact that the consumption vector is on the cone boundary is just by chance or whether there are economic forces at work to impose this condition. In the former case prices are undetermined and vary in a range. In the latter case the forces at work have an impact on prices which can be determined. The same argument is immediately applicable to an *n*-commodity economy: once again the no-arbitrage conditions determine a number of relations among prices lower than that which would be necessary to determine prices. Thus the economic forces which drive the consumption vector to the boundary may determine the further constraints which are able to complete the determination of prices. Otherwise prices are not fully determined and may vary in a range.

With these arguments in mind we can move on to analyze the last two regimes mentioned in the Classification Theorem. The third is clearly a case in which the consumption vector, which in our case is proportional to the first unit vector, is outside the cone: all commodities needed for the production of commodity 1 except commodity 1 itself are overproduced. The second regime mentioned in the Classification Theorem is clearly a case in which the consumption vector is on the boundary of the cone. In this range the no-arbitrage conditions are not able to fully determine prices; then in the long run an increment in ρ pushes down the prices of the inputs of commodity 1 pushing the rate of profit *r* up in such a way that the increment in ρ is exactly compensated and the growth rate is unchanged. Note that if the growth rate were pushed up, then the no-arbitrage conditions would impose prices proportional to $\mathbf{v}_{PF}$, whereas if it were pushed down, then the no-arbitrage conditions would impose zero prices for all inputs of commodity 1 apart from commodity 1 itself.

4.5. GENERALIZATIONS

The arguments developed in the previous section suggest that

- if there is joint production the second and third regime of the Classification Theorem do not need to be connected with negative growth rates,
- if more than one commodity is consumed (in a single production setting), then
  - the first regime of the Classification Theorem may also hold for values of the growth rate lower than $-\delta_x$,
  - if the proportion in which commodities are consumed depends on prices, then the second regime may determine a relationship between the rate of profit and the growth rate which may be different from a horizontal segment,
  - if there is continuous substitution in consumption, the third regime of the Classification Theorem may not exist.

Simple examples illustrating the above properties are easily constructed. In the appendixes to this chapter we provide such examples. Appendix A presents two examples involving joint production. The first example shows that the second regime of the Classification Theorem and part of the third can actually occur for positive growth rates. The second example shows that the first regime does not need to exist when joint production is involved.

Appendix B presents a number of examples involving two consumption commodities. These examples are related to the value assumed by a parameter. Three possibilities are envisaged. In all of them for high growth rates (i.e. low discount rates) the rate of profit and the prices are determined as in the Classification Theorem studied in FGS. However for low growth rates (i.e. high discount rates), the relationship between prices and quantities and growth and profit rates are very different from what was predicted by the theorem.

In general, the arguments developed in the previous section suggest that many of the results obtained in long-period models of Classical inspiration like those of von Neumann and Sraffa can at least partly be imported in the framework here presented. In particular the problem of choice of technique[11] and that of joint production appear to be easily handled. Similarly, some difficulties recognized in those models should have corresponding difficulties here. In particular we know that dropping Assumption [FGS4], or its general form [FGS3], may lead to difficulties. The analog of these difficulties in the present framework can be illustrated with an example. In Appendix C we present an example with two commodities: commodity 2 enters directly into the production of both commodities, whereas commodity 1 (which is the only commodity to be consumed) enters directly only into its own production. This simple model is analyzed to illustrate the difficulties that a decomposable matrix A can generate.[12] In this example, besides the

three regimes mentioned in the Classification Theorem, there is a fourth regime, in which commodity 2 behaves like a renewable resource which, if it is left to itself, grows at a rate higher than that of the consumption good. In order to produce commodity 1, producers pick it just as fruits were collected in the Garden of Eden: at no cost. This example can also be used to show what may happen when assumption [HP4] is dropped and, therefore, when storage is not part of the process of disposal but is effected in order to preserve the commodities. This point is clarified at the end of Appendix C.

4.6. CONCLUDING REMARKS

This chapter has investigated a number of problems which are absent in any single sector economy, but can be present in a multiple sector economy. This has been done with the help of a generalization of the multi-sector 'AK model' in continuous time which we analyzed in a previous paper. This analysis has also shown how this model is connected to the von Neumann–Sraffa–Morishima linear models investigated in the sixties and seventies.

APPENDIX A

Let (P) be the problem

$$V(\mathbf{s}^*) = \sup \int_0^\infty e^{-\rho t} \frac{c^{1-\sigma}}{1-\sigma} dt$$
$$\dot{\mathbf{s}}^T = \mathbf{x}^T\mathbf{B} - \delta_x \mathbf{x}^T\mathbf{A} - \delta_z(\mathbf{s}^T - \mathbf{x}^T\mathbf{A}) - \mathbf{c}^T$$
$$\mathbf{x}^T\mathbf{A} \le \mathbf{s}^T,\ \mathbf{x} \ge \mathbf{0},\ \mathbf{c} \ge \mathbf{0},\ \mathbf{s}(0) = \mathbf{s}^* > \mathbf{0}\ \ \textit{given.}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},\quad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = c\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is easily checked that the steady-state optimal solutions can be represented in the same three regimes referred to in the Classification Theorem. If $\delta_x < 1$, the second regime of the Classification Theorem and part of the third occur for positive values of the growth rate.

Let (P) be the same problem above, except that the input–output–matrices are now given by

$$\mathbf{A}=\begin{bmatrix}\frac{1}{4} & \frac{1}{2}\\ 0 & \frac{1}{2}\end{bmatrix} \text{ and } \mathbf{B}=\begin{bmatrix}2 & 2\\ 0 & 1\end{bmatrix}.$$

It is easily verified that the steady-state optimal solutions can be represented by two of the three regimes referred to in the Classification Theorem. The second process is inefficient with respect to the first process and therefore it is never operated. Hence the first regime of the Classification Theorem cannot exist.

APPENDIX B

Let (P) be the problem

$$V(\mathbf{s}^*)=\sup\int_0^\infty e^{-\rho t}\log(c_1^{1-\alpha}c_2^{\alpha})dt$$
$$\dot{\mathbf{s}}=\mathbf{x}-\mathbf{s}-\mathbf{c}$$
$$\mathbf{x}^T\mathbf{A}\le\mathbf{s}^T,\ \mathbf{x}\ge\mathbf{0},\ \mathbf{c}\ge\mathbf{0},\ \mathbf{s}(0)=\mathbf{s}^*\ge\mathbf{0}\ \ \textit{given}$$

where $\mathbf{A}=\begin{bmatrix}\frac{1}{8} & \frac{1}{4}\\ \frac{1}{8} & 0\end{bmatrix}$ and $\mathbf{c}=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$.

If $\alpha = 0$ or $\alpha = 1$, there is a single consumption commodity. Hence, let $0 < \alpha < 1$. The instantaneous Cobb–Douglas utility function determines consumption share, in value, as constant and depending only on α: along any optimal path

$$\frac{c_1v_1}{c_2v_2}=\frac{1-\alpha}{\alpha}$$

and, as a consequence, both prices and both consumption levels are positive. Hence in a steady-state optimal solution

$$c_1=x_1-(1+g)s_1>0,\ c_2=x_2-(1+g)s_2>0$$

Let us partition all possible cases on the basis of operated processes. If both processes are operated, then

$$r=3,\ \frac{v_1}{v_2}=2\ \ ,\ \ \frac{c_1}{c_2}=\frac{1-\alpha}{2\alpha},$$

$$x_1=\frac{8(4-3\alpha+\alpha g)}{(1-\alpha)(27-6g-g^2)}c_1\ \ ,\ \ x_2=\frac{1+6\alpha+g-2\alpha g}{(1-\alpha)(27-6g-g^2)}c_1.$$

There are two critical values for the parameter α: 1/3, and 1/2. For $0<\alpha<1/3$ vector $\mathbf{x}$ is positive if and only if $g<[(3\alpha-4)/\alpha]$ or $[(1+6\alpha)/(2\alpha-1)]<g<3$. For $\alpha = 1/3$ vector $\mathbf{x}$ is positive for $g < 3$. For $1/3<\alpha<1/2$ vector $\mathbf{x}$ is positive if and only if $g<[(1+6\alpha)/(2\alpha-1)]$ or $[(3\alpha-4)/\alpha]<g<3$. For $1/2\le\alpha<1$ vector $\mathbf{x}$ is positive if and only if $[(3\alpha-4)/\alpha]<g<3$. If only the first process is operated, then

$$\frac{v_1}{v_2}=\frac{2(1+r)}{7-r}\ \ ,\ \ \frac{c_1}{c_2}=\frac{g-7}{2(1+g)}\ \ ,\ 3<r<7\ \text{ and }\ g<-1.$$

Hence $g=[7+(8\alpha-1)r/(8\alpha-7)+r]$ and $\alpha < 1/2$. If only the second process is operated, then

$$\frac{v_1}{v_2}=\frac{8}{1+r}\ ,\ \frac{c_1}{c_2}=-\frac{(1+g)}{8}\ ,\ r>3\ \text{ and }\ g\le-1.$$

Hence $g=-[1+(1-\alpha)r/\alpha]$. Figures 4.4 provide the relationship between g and r for α‘s in one of the three relevant ranges: $0<\alpha<1/3$, $1/3<\alpha<1/2$, $1/2\le\alpha<1$, respectively.

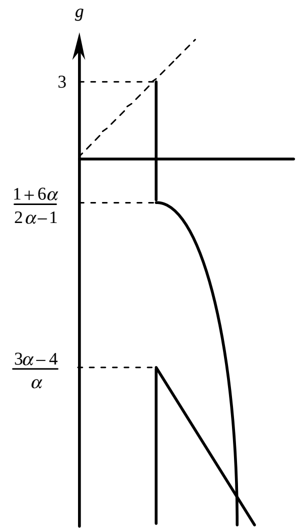


Figure 4.4a

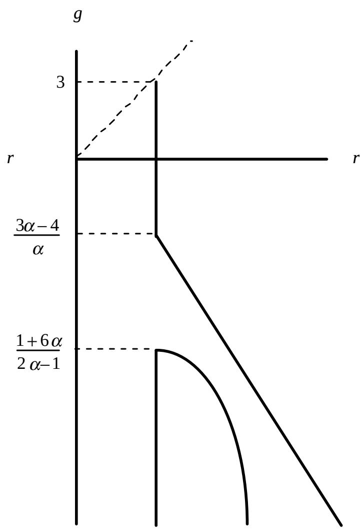
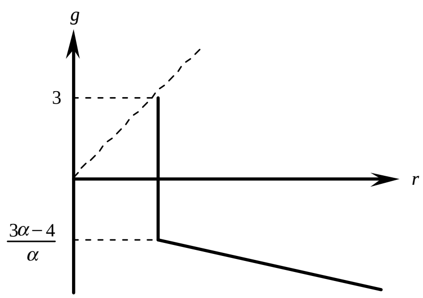


Figure 4.4b



Figures 4.4c

APPENDIX C

Let (P) be the problem

$$V(\mathbf{s}^*) = \sup \int_0^\infty e^{-\rho t} \frac{c^{1-\sigma}}{1-\sigma} dt$$
$$\dot{\mathbf{s}}^T = \mathbf{x}^T \mathbf{I} - \delta_x \mathbf{x}^T \mathbf{A} - \delta_z (\mathbf{s}^T - \mathbf{x}^T \mathbf{A}) - \mathbf{c}^T$$
$$\mathbf{x}^T \mathbf{A} \leq \mathbf{s}^T,\ \mathbf{x} \geq \mathbf{0},\ \mathbf{c} \geq \mathbf{0},\ \mathrm{s}(0) = \mathbf{s}^* > \mathbf{0}\ \ \textit{given}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix},\ a_{11} < a_{22} \text{ and } \mathbf{c} = c\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is easily verified that the steady state optimal solutions can be represented in four regimes which are depicted in Figure 4.5. Three regimes are the same mentioned in the Classification Theorem. The fourth regime is characterized by $r = a_{11}^{-1} - \delta_x$ and $-\delta_x < g < a_{22}^{-1} - \delta_x$. In this regime commodity 2 is (over)produced and its price equals 0. The growth rate cannot be equal or larger than $a_{22}^{-1} - \delta_x$ because otherwise the intensity of operation of the process producing commodity 1 would be nought or negative (and therefore consumption would be nought or negative).

Process 2 in the example can be interpreted as a storage process for commodity 2 with a *negative* rate of decay. The existence of the fourth regime, however, does not depend on the sign of the rate of decay, but on the fact that $a_{22}^{-1} - \delta_x > -\delta_x$. It is clear therefore that, even if matrix **A** is indecomposable, something similar to the fourth regime in this example comes into existence whenever assumption [HP4] does not hold.

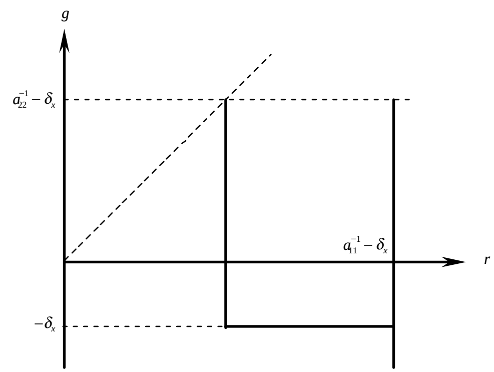


Figure 4.5

NOTES

1. Champernowne (1945, p. 12) used this expression for the von Neumann (1945) model. In the following we will borrow other expressions from this paper by Champernowne to emphasize some similarities with the von Neumann model and the literature devoted to it.
2. Some theorists even came to the conclusion that unbounded growth is more an assumption about the linearity of the technology than a result of the models; see for instance Romer (1990, p. S84).
3. This is not the usual interpretation found in the recent literature. On the contrary it has been suggested that the AK model 'becomes more plausible if we think of K in a broad sense to include human capital' (Barro and Sala-i-Martin, 1995, p. 39). Such two interpretations are not so different as they seem at first sight. We will come back on this in footnote 4. It also possible to assume that the necessary subsistence of workers (or the inputs to produce human capital) are included in the A of the AK model, whereas another part of wages, which exceeds the necessaries, are subject to the choice of the 'representative agent', who, in this case, does not need to be anymore a 'representative capitalist'.
4. In this description of technology labour was not explicitly considered. On the contrary a process can be represented as

$$\mathbf{a}_j^T \oplus l_j \rightarrow \mathbf{b}_j^T,$$

where l_j refers to labour input. Then there are two ways to obtain the simbolism used in the text, in the assumption that the real wage rate per unit of labour is defined by the vector w. Let A be the usual material input matrix used in input–output analysis and let l be the input vector of (simple) labour. Then the matrix A in the text can be seen either as $A = \hat{A} + lw^T$ or as

$$\mathbf{A} = \begin{bmatrix} \hat{\mathbf{A}} & 1 \\ \mathbf{w} & 0 \end{bmatrix}.$$

In the latter alternative the last process referes to production' of 'labour' or 'human capital'.

5. The use of iso-elastic utility functions goes back to Ramsey (1928) who studied this ‘interesting special case’ in section II of his 1928 paper.
6. Cases with multiple consumption goods are considered by Gozzi and Freni (2001).
7. We note that the utility function u is unbounded above or below or, in the log case, both above and below. Moreover, the boundedness of feasible paths is not assumed. This generates a non-trivial existence problem for (P) which is usually solved by the introduction of an existence condition linking the technology with the preferences (see for example McFadden, 1967).
8. An example can clarify the issue. Let

$$\rho = 1/3\ ,\ u(\mathbf{c}(t)) = 2c_1^{1/2}\ ,\ \mathbf{B} = \mathbf{I}\ ,\ \delta_x = \delta_z = 1/2\ ,\ \mathbf{A} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix},\ \hat{\mathbf{c}} = c\,\mathbf{e}_1 \text{ and } \mathbf{s}^* = \begin{bmatrix} \frac{42}{31} \\ \frac{72}{31} \\ \frac{72}{31} \end{bmatrix}$$

in the (P) problem. It is easily checked that

$$\mathbf{s} = \mathbf{s}^* e^{\frac{1}{3}t}\ ,\ \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} e^{-\frac{1}{6}t} + h \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} e^{-\frac{7}{6}t}\ ,\ c = e^{\frac{1}{3}t}\ ,\ \mathbf{x} = \begin{bmatrix} \frac{66}{31} \\ \frac{60}{31} \\ \frac{60}{31} \end{bmatrix} e^{\frac{1}{3}t}\ ,\ \mathbf{q} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} e^{-\frac{1}{6}t} + 2h \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} e^{-\frac{7}{6}t}$$

for $-1/2 < h < 1/2$ is a steady-state optimal solution to problem (P) supported by a competitive rental-price path, so that despite the fact that the quantity side grows at rate 1/3, the relative prices and relative rentals do not need to be constant (they are so if and only if h = 0).
9. This formula is clearly a reminiscence of Fisher formula, when the own rate of return is interpreted as a ‘real’ rate of profit, the discount rate as a ‘nominal’ rate, and the fraction in the RHS as an inflation rate
10. Evans et al. (1998) refer to this relationship as a ‘technological’ relationship whereas they refer to equation (5) as a ‘preferences’ relationship. However, as shown in the example of Appendix C, the former relationship depends not only on technology, but also on preferences concerning consumption at a given moment in time.
11. It is easily checked that the Classification Theorem can be generalized to allow the existence of several processes for the production of each commodity (even a continuous number). If Assumptions [FGS1], [FGS3], [FGS5], [FGS6] hold and if it is further assumed that each process producing commodity 1 uses commmodity 1 as an input, the Classification Theorem holds with the following differences. In the first regime the rate of profit is $\lambda - \delta_x$, where λ is the minimum of all the Perron–Frobenius eigenvalues of possible input square matrices which can be obtained by peaking up for each commodity a process producing it; in the third regime the rate of profit is $a^{-1} - \delta_x$, where a is the minimum of all coefficients relative to inputs of commodity 1 used in the production of commodity 1.
12. The case of decomposable matrices is not uncommon in the new growth literature: the model by Lucas (1988), for instance, is of this type (for a similar remark, see also McKenzie, 1998, p. 11).

5. Income distribution and consumption patterns in a ‘classical’ growth model

Davide Fiaschi and Rodolfo Signorino

5.1. INTRODUCTION

Historians of the Industrial Revolution have not failed to study the role played by demand factors in the process of industrialization of an agricultural economy. Landes (1969, ch. II) emphasizes the relation between income distribution, consumption patterns and the growth of manufactures in the eighteenth century England. In his view, the middle classes flourished thanks to favourable income and wealth distribution. The typical consumption pattern of these classes consisted of commodities manufactured using mass production techniques with a high capital/labour ratio. Moreover, English farmers were used to eating a superior kind of food, such as white bread, and to spending a smaller share of their income on food than their Continental counterparts. Thus English farmers had more money to spend on non-agricultural commodities. Such a consumption pattern is, for Landes, one of the key elements which favoured the industrialization of the English economy.

Landes’ point of view is all the more intriguing when confronted with that endorsed by some leading British classical economists who were direct witnesses of the historical facts that he studied many decades later. We refer in particular to Adam Smith and Thomas Robert Malthus (see Rosenberg, 1968; Brewer, 1998; and Marshall, 2000). In Book III of *The Wealth of Nations* (hereafter WN), Smith reconstructs the progress of wealth in Europe from the fall of the Roman Empire and highlights the role played by landowners’ consumption patterns. In a stationary agricultural economy an increasing taste for ‘luxuries’, usually imported from abroad, provides landowners with a powerful stimulus to modify their routine economic behaviour. Landowners are willing to change the lease conditions to their tenants in order to allow the latter to implement more efficient agricultural techniques and, consequently, to pay higher rents. Thus productivity in the key sector of the economy, agriculture, increases. Landowners’ increasing expenditure on luxuries makes domestic production of these commodities profitable. According to Smith, in fact, ‘finer manufactures’ were introduced into agricultural economies either through the gradual refinement of domestic primitive manufactures or through the imitation of foreign manufactures (see *infra* Section 5.2).

Malthus’ analysis of the process of growth in an industrialized economy is found in Book II of his *Principles of Political Economy* (hereafter PPE). According to Malthus, the basic obstacle which may slow down growth in an industrialized economy is the lack of ‘an adequate stimulus to the continued increase of wealth’ (Malthus, 1986, p. 288). This stimulus consists in an adequate level of ‘effectual demand’ mainly determined by income distribution and the structure of property rights. For Malthus a wide class of relatively well-off farmers is able to generate a level of expenditure much higher than that generated by few large landowners (when land ownership is too highly concentrated) or by a multitude of poor peasants (when land ownership is too fragmented). By the same token, the extent of internal and external trade and the level of consumption from Smithian ‘unproductive labourers’ need not be too low for effectual demand and industrial production to grow *pari passu* (see *infra* Section 5.3). Moreover, as real wages increase, workers may develop a taste for manufactured commodities, usually referred to as ‘conveniences’ or ‘comforts’, which may induce them to control their fertility and resist the temptation of ‘indolence or love of ease’. As Gilbert points out:

> When men are seen exercising a free choice not to marry at the first opportunity, it becomes more difficult to view them crudely as mere food-consumers and children-producers (Gilbert, 1980, p. 90).

In this scenario, workers’ rational decisions concerning their fertility and consumption basket may become a crucial variable affecting the long-run growth performance of an industrialized economy. By contrast, analysis of demand factors and, in particular, analysis of the relationship between income distribution, consumption patterns and growth is not high on the contemporary research agenda. Among the few exceptions it is possible to mention Laitner (2000), Zweimuller (2000) and Kongsamut et al. (2001). Laitner analyzes the structural change of an economy with two goods and non-homothetic preferences, focusing on the wealth effects involved by such a transition. Zweimuller is interested in the relationship between demand composition and innovation and investigates the properties of the balanced growth path. Finally, Kongsamut et al. study a multisectoral economy which shows structural change and whose income grows at a constant rate.

The basic aim of our paper is to present a model in which the take-off of an agricultural economy as well as the long-run growth of an industrialized economy depends on income distribution and consumption patterns. Our model is an extension of Murphy et al. (1989), up till now unduly neglected in contemporary literature. The main differences (crucial to our findings) are the following: (*i*) agricultural productivity is a function of land distribution and of the availability of industrial goods; (*ii*) workers’ population is endogenous; and (*iii*) wage-earners may consume both agricultural goods and industrial goods.

Besides economic historians and classical economists, another important source of inspiration for our paper has been the work of Luigi Pasinetti (1981, 1993) on structural economic dynamics. To put it in a nutshell, according to Pasinetti, economic growth implies structural dynamics: demand composition greatly changes as GDP increases (the so-called Engel’s Law) and technical progress seldom displays the same rate in all productive sectors. The economic structure of growing economies is thus bound to change. Hence, to be empirically relevant, multisectoral models of economic growth should dispose of the assumption of proportional growth. Moreover, Pasinetti depicts technical progress and the evolution of the patterns of demand, the two great dynamic forces driving growth in real world economies, as two intertwined phenomena:

> Since increases in per capita income necessarily imply non-proportional expansion of demand, and since technical progress leads to increases in per capita incomes, the introduction of technical progress in any dynamic economic investigation necessarily implies a non-proportional expansion of demand (Pasinetti, 1981, p. 70).

Conversely,

> No commodity, whatever ingenious technique it may require, can be successfully produced if its (real or imagined) utility for the consumers is not sufficient to justify its cost: it would remain unsold. The relevance itself of technical progress depends on potential demand: an increase of productivity, however large it may be, loses much or even all of its meaning, if it takes place in the productive process of a commodity for which demand can only be small or negligible. This means that any investigation into technical progress must necessarily imply some hypotheses … on the evolution of consumers’ preferences as income increases (Ibidem, pp. 68–9).

Pasinetti focuses his analysis on the process of growth of industrial economies. We make use of his intuitions on the importance of Engel’s Law and increases of productivity also to study the development process of an agricultural economy.

The paper has seven sections. Sections 5.2 and 5.3 briefly recall Smith's and Malthus' points of view on the relation between property rights distribution, consumption expenditure, development and growth. Section 5.4 outlines the formal model whose equilibrium conditions are investigated in Section 5.5. In Section 5.6 we use our model to formalize three interesting stages in the development process of an economy. Finally, Section 5.7 draws some concluding remarks.

5.2. LANDOWNERS' CONSUMPTION AND THE TAKE-OFF OF A STATIONARY AGRICULTURAL ECONOMY

According to Smith, 'no large country ... ever did or could subsist without some sort of manufactures being carried on in it. [...] This is even more universally the case in those poor countries which are commonly said to have no manufactures' (WN, III.iii.17). Smith's examples of primitive manufactures carried on in all countries are clothing and housing. Of course, the presence of such manufactures does not imply that the economy is an industrial one. Thus a stationary agricultural economy may be defined as an economy where there are no manufactures of the industrial type and where rent, earned by landowners (the sovereign, the landlords and the clergy), is spent on the consumption of personal services or of commodities produced by foreign 'finer manufactures' and imported from abroad.

Rent may be also partially saved and hoarded.[1] It goes without saying that in a stationary agricultural economy rent is the only kind of income which is not tied to the requirement of (re)production or of subsistence. Rent absorbs the whole surplus produced in the economy.

It is to be stressed that landowners' consumption behaviour is not explained by Smith in subjectivistic terms, e.g. as the preference for a certain kind of commodities. Smith's explanation runs in objectivistic terms since it involves the absence in the economy under scrutiny of alternative commodities to consume. Thus landowners' consumption behaviour is caused by the narrowness of the commodity space at their disposal, which is a very important point in relation to the possibility of take-off of the economy. Moreover, the absence of finer commodities to consume implies the absence of incentives to improve the organization of agricultural production. As noted by Brewer, given known technology, land resources are under-employed and technological innovation is neglected:

> agriculture had been ... under-performing because of indolence, caused by a lack of attractive manufactures. A taste for 'luxury', and an opportunity to gratify it, provides the incentives which are the key to economic development (Brewer, 1998, p. 81).

A widespread taste for finer commodities, in fact, creates incentives for landowners to modify dramatically their overall economic behaviour. On the one hand, landowners are led to cut their expenditure on personal services and hospitality and to increase their expenditure on commodities; on the other hand they are willing to change the lease conditions to their tenants in order to allow them to improve agricultural technique and thus to pay higher rents:

> Farms were enlarged, and the occupiers of land, notwithstanding the complaints of depopulation, reduced to the number necessary for cultivating it, according to the imperfect state of cultivation and improvement in those times. By the removal of the unnecessary mouths, and by exacting from the farmer the full value of the farm, a greater surplus ... was obtained for the proprietor, which the merchants and manufacturers soon furnished him with a method of spending upon his own person in the same manner as he had done the rest. The same cause continuing to operate, he was desirous to raise his rents above what his lands, in the actual state of their improvement, could afford. His tenants could agree to this upon one condition only, that they should be secured in their possession, for such a term of years as might give them time to recover with profit whatever they should lay out in the further improvement of the land. The expensive vanity of the landlord made him willing to accept this condition; and hence the origin of long leases (WN, III.iv.13).

Thus, according to Smith, the implementation and development of manufactures deriving from a widespread taste for finer commodities involve a series of interesting economic phenomena: the widening of the commodity space leads to a significant improvement of the efficiency of agricultural production and to an increase of both national income and consumption. Rosenberg summarizes Smith's point of view in this regard:

> The expansion in the range of alternatives for the disposition of the economic surplus had the immediate effects of 1) shifting the composition of consumer expenditure flows away from services and towards goods; 2) shifting upward the consumption functions of large property owners, who previously lived within their incomes because of the limited scope afforded for the exercise of personal vanity; and 3) the strength of the desire for these new goods provided a motive for efficient cultivation which was previously lacking. The increased incentive provided by the availability of new goods led to the elimination of known inefficiencies which had previously been tolerated and to legal and institutional changes which, by strengthening economic incentives, Smith regarded as indispensable to sustained economic growth (Rosenberg, 1968, p. 368).

In a stationary agricultural economy the possibility of take-off is strictly tied to the creation of conditions favourable to the implementation and development of manufactures. In the economy under scrutiny, the traditional sectors, such as agriculture and personal services, have plenty of physical resources which are idle or under-employed and which may be diverted to the new, growing sector, manufactures. Yet, the availability of idle physical resources is not sufficient for take-off. The presence is required of a class of agents within the economy who gain profits from the refinement of domestic primitive manufactures. This class gains benefits from the fact that the taste for finer commodities has become so general as to occasion a considerable demand for them. In the case of an economy characterized by a wide external trade, profits may also derive from the domestic production of those finer commodities previously imported from abroad.

5.3. CONSUMPTION PATTERNS AND INCOME DISTRIBUTION IN AN INDUSTRIALIZED ECONOMY

Once industrialization starts, the economic problem is constituted by the persistence of growth, that is, by the cumulative processes which may sustain or choke the expansion of manufactures. Smith apparently privileges supply side factors in the capital and labour markets. Right at the beginning of WN Smith declares that in each country per capita income is not regulated by natural factors such as 'soil, climate, or extent of territory' but by the productivity of its labourers, namely 'first, by the skill, dexterity, and judgment with which its labour is generally applied; and, secondly, by the proportion between the number of those who are employed in useful labour, and that of those who are not so employed'. Since only part of total labour is applied to productive activities, the rate of growth of a country is greatly affected by the factors which determine the share of 'productive labourers' and its dynamics. These factors are basically saving decisions: it is 'parsimony', in fact, which determines the 'funds destined for the maintenance of productive labour' (WN, II.iii.14).

Yet, Smith did not neglect the demand side in his analysis of economic growth. In a growing economy the leading sector is the manufacturing sector: manufactures are characterized by considerable scale economies and a high rate of technical progress because they offer broader scope than agriculture for the process of division and specialization of labour. As is well known, for Smith, the pace of this process is basically determined by demand factors (WN, I.iii). As Young remarked 152 years after Smith: 'it would be wasteful to make a hammer to drive a single nail' (Young, 1928, p. 530).

Also Ricardo focuses mainly on supply factors as the main determinants of growth. For Ricardo, limits to growth are basically to be located in the supply conditions of a crucial factor of production, land, whose quantity and quality are assumed to be given and invariant. The logical chain underlying Ricardo's argument may be briefly reconstructed as follows. In the long run the rate of real wages is at its historically determined level of subsistence. Workers consume almost exclusively agricultural products (usually referred to as 'corn'). Thus in the long run the money price of 'corn' regulates the rate of money wages. Since the 'natural' or long-run normal price of each commodity is determined by its productive conditions, the normal price of 'corn' is regulated by the state of cultivation. Thus diminishing returns in agriculture provoke an increasing 'price of labour' for industrial entrepreneurs (the rate of money wages increases to compensate the rising price of 'corn'). Provided that 1) the rate of wages and the rate of profits move in opposite directions and that 2) the rate of profits and the rate of capital accumulation move in the same direction, an increasing population provokes an increasing demand for 'corn', a rising price of 'corn' (as soon as agriculture enters its diminishing returns stage), a rising rate of wages, a falling rate of profits and a falling rate of capital accumulation. Population growth and capital accumulation cease as soon as the rate of wages and the rate of profits reach their 'natural' levels. If technical progress is unable to counteract the action of diminishing returns in agriculture, the process of economic growth inevitably comes to a halt.

It is not difficult to find quotations within Ricardo's texts which point to a supremacy of supply factors in the analysis of growth. Perhaps the most explicit is the following:

> Profits do not necessarily fall with the increase of the quantity of capital because the demand for capital is infinite and is governed by the same law as population itself. They are both checked by the rise in the price of food, and the consequent increase in the price of labour. If there were no such rise, what could prevent population and capital from increasing without limit? (Ricardo, 1951–73, vol. VI, p. 301).

Yet, things are not that easy. One of the crucial assumptions in the chain of reasonings sketched above is that workers consume (almost) only 'corn', that is, a basket of commodities produced in the agricultural sector and thus subject to the law of diminishing return. What would happen if workers reacted to an increasing real income with a reduction of their consumption of 'corn', with a control of their fertility and with an increase in their consumption of 'conveniences', that is, manufactured commodities produced under a regime of increasing returns? In short, what would happen if workers' consumption baskets abided by Engel's Law, one of the most certain empirical regularities in economics? In his paper on the mathematical formulation of the Ricardian system, Pasinetti (1960) overtly denounces ‘the crudeness of Ricardo’s assumptions’:

> The economic theory of demand had not yet been developed, at [Ricardo’s] time, and there is no question of substitution among wage goods in the Ricardian model. The *natural* wage-rate is represented by a fixed *basket of goods*, to be accepted as given by factors lying outside economic investigation (Pasinetti, 1960, p. 90, Pasinetti’s emphasis).

Ricardo himself was aware that to assume the normal consumption basket of workers as consisting only of ‘corn’ was too crude an assumption even for his times. He plainly acknowledges that workers’ standards of living are constantly rising:

> Many of the conveniences now enjoyed in an English cottage, would have been thought luxuries at an earlier period of our history (Ricardo, 1951–73, vol. I, p. 97).

He also acknowledges that in a growing economy the movement of relative prices favours an increasing consumption of manufactured commodities:

> From manufactured commodities always falling, and raw produce always rising, with the progress of society, such a disproportion in their relative value is at length created, that in rich countries a labourer, by the sacrifice of a very small quantity only of his food, is able to provide liberally for all his other wants (Ibidem).

Finally, Ricardo is perfectly aware that a widespread taste for conveniences among workers is the surest remedy against the evils deriving from the action of the so-called Malthusian Law of population:

> The friends of humanity cannot but wish that in all countries the labouring classes should have a taste for comforts and enjoyments, and that they should be stimulated by all legal means in their exertions to procure them (Ricardo, 1951–73, vol. I, p. 100).

Unfortunately, Ricardo did not develop these interesting insights any further. By contrast, in Malthusian economics the rate of profits and the rate of capital accumulation may decline even if top quality land is still available. Limits to growth for Malthus are basically to be located on the demand side. Malthus concedes to Ricardo that the scarcity of land of first quality is the limiting principle of profits; but he claims that the principle which actually regulates the rate of profits is ‘the varying value of the produce of the same quantity of labour occasioned by the accidental or ordinary state of the demand and supply, by which a greater or smaller proportion of that produce

falls to the share of the labourers employed’ (Malthus, 1986, p. 219).[2] Malthus’ argument is that a low level of ‘effectual demand’ for manufactured commodities may depress the rate of profits and the rate of capital accumulation long before the exhaustion of fertile lands:

> But it appears to me perfectly clear in theory, and universally confirmed by experience, that the employment of capital may, and in fact often does, find a limit, long before there is any real difficulty in procuring the means of subsistence; and that both capital and population may be at the same time, and for a period of considerable length, redundant, compared with the effectual demand for produce (Malthus, 1986, p. 321).

For Malthus, a low level of ‘effectual demand’ may derive from three sources: (*i*) a ‘most unequal and vicious’ distribution of land resources, (*ii*) barriers to internal and external trade and (*iii*) an insufficient amount of consumption from Smithian unproductive labourers. Malthus explicitly considers the distribution of land property as one of the main determinants of the level of ‘effectual demand’. In his view

> a very large proprietor, surrounded by very poor peasants, presents a distribution of property most unfavourable to effectual demand. [...] Thirty or forty proprietors, with incomes answering to between one thousand and five thousand a year, would create a much more effectual demand for the necessaries, conveniences, and luxuries of life, than a single proprietor possessing a hundred thousand a year (Malthus, 1986, pp. 298–9).

Yet Malthus was not unaware that consumption expenditure may be depressed by going too far in the redistribution of land property. In the final part of Chapter VII, in fact, he discusses the economic consequences of the abolition of the right of primogeniture in England and France: a possible long-run outcome of this abolition may be an excessive fragmentation of land property which may have a negative effect on the development of a country.

The following Chapter VIII is devoted to scrutinizing the relationship between the extent of internal and external trade and domestic prosperity.[3] According to Malthus, ‘no country with a very confined market, internal as well as external, has ever been able to accumulate a large capital, because such a market prevents the formation of those wants and tastes, and that desire to consume, which are absolutely necessary to keep up the market prices of commodities, and prevent the fall of profits’ (Malthus, 1986, p. 309).

Consumption expenditure out of wages may play for Malthus an important role to sustain long-run growth. In Chapter IV of PPE Malthus admits that growing real wages may induce workers to modify their concept of subsistence not only from a quantitative but also from a qualitative point of view. In rapidly growing economies, real wages and, consequently, workers' power of purchasing manufactured commodities or 'conveniences' usually increase. Since the consumption of subsistence goods, usually referred to by classical authors as 'necessaries' or more simply 'corn', is proportional to the number of children to rear, a family of growing size obliges workers to reduce the share of their income devoted to the consumption of conveniences. Thus if workers develop a taste for conveniences, then they would probably control their fertility and keep their labour supply at least constant in the face of increasing real wages:

> From high real wages, or the power of commanding a large portion of the necessaries of life, two very different results may follow: one, that of a rapid increase of population, in which case the high wages are chiefly spent in the maintenance of large and frequent families; and the other, that of a decided improvement in the modes of subsistence, and the conveniences and comforts enjoyed, without a proportionate acceleration in the rate of increase (Malthus, 1986, p. 183).

In this second scenario rising real wages provide an important source of effectual demand for commodities produced by manufactures.

5.4. THE BASIC MODEL

In this section we present an extension of Murphy et al. (1989) whose basic features are the following. The economy has an agricultural sector, a manufacturing sector and a personal services sector. The former produces a homogeneous good, 'food', by means of a decreasing returns technology using land and labour. The manufacturing sector produces a continuum of goods by means of an increasing returns technology using labour: we call these goods 'industrial goods'. We assume that the very same goods may be produced by a constant returns technology using labour in the personal services sector: we call these goods 'personal services'. The economy is populated by a continuum of agents of measure L. We assume that each agent devotes his/her first z units of income to food consumption; all the remaining income (if any) is devoted to non-agricultural goods x. The latter are ranked according to an index q and the marginal utility of good q is decreasing in q. Every non-agricultural good is such that a consumption lower or greater than one unit does not produce any utility. More formally, let $X=\{x(q)\in\{0,1\}\ with\ q\in[0,+\infty)\}$ be the consumption set of non-agricultural goods. The assumptions made on preferences imply that every

Figure 5.1 – Engel's curves for food and non-agricultural goods

Thus the consumption pattern of each agent depends on his/her personal income; as long as income is lower than a certain threshold, z, total income is devoted to food consumption. Rich agents can consume both food and non-agricultural goods; richer agents consume a wider range of goods than poorer agents. Agents care about current consumption; saving and investment in our model are carried on in the same period and there is no proper accumulation of capital stock as in standard growth models. As shown below, we assume that investment takes the form of a payment of a fixed cost in terms of wages and has a depreciation rate equal to 1.

Food is produced by land and labour. We assume a fixed coefficient production function and decreasing returns to scale. Moreover, we assume that the supply of land is overabundant in relation to labour supply. Employment in agriculture is thus the only factor which determines the aggregate production of food. We consider the price of food as numeraire, so that the following equality holds:

$$F(L_F, a_F) = R + w_F L_F \tag{1}$$

where R is the rent, F the production function in the food sector, with $\partial F/\partial L_F > 0$ and $\partial^2 F/\partial L_F{}^2 < 0$, a_F is a productivity parameter $(\partial F/\partial a_F > 0)$ and w_F is the rate of wages in the food sector. We assume that w_F negatively depends on L_F, while R positively depends on L_F:

$$w_F = w(L_F, a_F) \tag{2}$$

$$R_F = R(L_F, a_F) \tag{3}$$

where $\partial w/\partial L_F < 0$, $\partial w/\partial a_F > 0$, $\partial R/\partial L_F > 0$ and $\partial R/\partial a_F > 0$.[5] We stress that the assumption $\partial w/\partial L_F < 0$ alone does not imply by itself that $\partial R/\partial L_F > 0$.

Every good can be produced by two technologies. If a good is produced by the increasing returns technology we consider it an industrial good or a convenience; if the good is produced by the constant returns technology we consider it a personal service or a luxury good.[6] In particular, a_L units of labour are necessary to produce one unit of good when the constant returns technology is adopted, while $a_M < a_L$ units of labour are necessary when the increasing returns technology is adopted. Yet, in order to use the latter technology firms have to pay a fixed investment equal to *C* units of labour. This cost can be considered an *R&D* activity which allows firms to discover a new method of production and provides them with a monopoly power.[7] By contrast, luxury goods are sold in competitive markets.

Though agents' endowments of land and capital are different, for simplicity we assume that the share of land owned by each agent is the same as his/her share of firms. This implies that there is no difference between capitalists and landowners and that rents and profits distributions are the same. Once agents are ranked according to their shares in increasing order, s_i is the share of agent *i* and $G(s_i)$ the cumulative distribution of *s*. We assume that a large part of population owns no property rights on land and/or firms. Let *L* be the total population and $N = L(1 - G(\underline{s}))$ be the number of shareholders, where $\underline{s} = s_{L-N}$ is the minimum positive share: $L - N$ is the number of agents having nothing but their labour. The generic agent's income is given by $y = w + s(R + \Pi)$, where $s \in \{0, [\underline{s}, \infty)\}$ and Π is the aggregate profit. We assume that every shareholder can buy at least *z* units of food:

$$y = w + \underline{s}(R + \Pi) > z \tag{4}$$

The value of *N* and the shareholders' distribution are crucial in order to determine the extent of demand of conveniences if $w < z$, that is, if wage-earners can buy only food; otherwise the demand for conveniences depends on the size of population *L*.

5.5. EQUILIBRIUM ANALYSIS

In this section we analyze the equilibrium characteristics of our model, given technology. Murphy et al. (1989b) consider labour in agriculture and labour in manufactures as imperfect substitutes. According to historical observation, they suppose that $w_M > w_F$ because of an indirect cost borne by labourers to work in a factory. Hence, if $w_M > z > w_F$ then a flow of labour from agriculture to manufactures implies an increase of demand for industrial goods. Nonetheless, for the sake of simplicity, we assume that labourers can costlessly move across sectors and that every agent supplies one unit of labour inelastically. This implies that wages in all sectors are equal:

$$w_F = w_M = w_L = w \tag{5}$$

The equilibrium condition in agriculture determines the level of employment in agriculture, L_F, together with conditions (2) and (5). In equilibrium the following relationship must hold:

$$\min\{w, z\}(L - N) + Nz = F(L_F, a_F) \tag{6}$$

From the above equality it is possible to derive the equilibrium employment in agriculture (according to equation (2) *w* is a function of L_F). Finally since the following relationship must hold:

$$wL_F + R = \min\{w, z\}(L - N) + Nz$$

the level of rents in equilibrium is:

$$R = \min\{w, z\}(L - N) + Nz - wL_F \tag{7}$$

In equilibrium the extent of industrialization is determined by the demand for conveniences, given technology and prices. Murphy et al. (1989) show that the equilibrium prices of all non-agricultural goods are the same and equal to the price which obtains when the good is produced by the constant returns to scale technology. In particular, since the markets for luxuries are competitive then the price of a luxury good is set at $a_L w$ and the profit for a monopolist is given by

$$\pi = (a_L w - a_M w)O - Cw$$

where *O* is the market output. A good is produced if (expected) profits are positive, that is, if $O \geq C/(a_L - a_M)$. Therefore, the minimum quantity which makes it profitable to produce a good is O^*, given by

Figure 5.2 – Income distribution and consumption pattern

To complete the description of equilibrium it is necessary to calculate total profits. The latter are equal to

$$\Pi = \frac{a_L w - a_m w}{a_L w}\left\{(L-N)\max\{w-z,0\} + (R+\Pi)\int_{\underline{s}}^{s^*} s\,dG(s)\right\} +$$
$$+\frac{a_L w - a_m w}{a_L w}\left\{(N-N^*)(w-z) + N^* s^* (R+\Pi) + N^*(w-z)\right\} - CwQ^*$$

where the first member in brackets is the demand (in nominal terms) from labourers (residual income from food); the sum of the second and third members is the demand from middle class shareholders and the sum of the fourth and fifth members is the demand from the richest agents; ($a_L w - a_M w$) is the difference between the price and the average (and marginal) cost of production and finally CwQ^* represents all the fixed costs paid in the economy.

Substitution from (10) yields:

$$\Pi = \eta\left\{(L-N)\max\{w-z,0\} + (R+\Pi)S^M + (N-N^*)(w-z)\right\}$$

where

$$\eta = \frac{a_L w - a_M w}{a_L w}$$

is the mark-up in the monopolist markets of industrial goods and

$$S^M = \int_{\underline{s}}^{s^*} s\,dG(s)$$

is the share of profits and rents owned by the middle class. Note that, given $N^* \leq N$, if $w > z$ then profits are always positive. This also holds if $w < z$ since $S^M(R+\Pi) > (N-N^*)(z-w)$ by assumption (4).[8] Therefore $N^* \leq N$ is a necessary and sufficient condition for the existence of a manufacturing sector. Rearranging, we obtain:

$$\Pi = \frac{\eta\left\{(L-N)\max\{w-z,0\} + RS^M + (N-N^*)(w-z)\right\}}{1-\eta S^M} \qquad (11)$$

Aggregate profits are a positive function of aggregate rents *R*, of the level of wages *w*, of the aggregate share owned by the middle class S^M and of mark-up η. Finally, it is interesting to calculate employment in the manufacturing sector:

$$L_M = \frac{(L-N)\max\{w-z,0\}+\left(S^M+N^*s^*\right)(R+\Pi)+N(w-z)}{a_L w}$$

which positively depends on $S^M + N^* s^*$ (the total share of profits plus rent spent in conveniences) and negatively on a_L. A manufacturing sector cannot arise if (*i*) $w > z$ and $O^* > L$ or if (*ii*) $w < z$ and $O^* > N$. If a manufacturing sector cannot arise, then only the first three equations and the last are the relevant ones, with $L_M = 0$.

5.6. EXTENSIONS OF THE BASIC MODEL

In this section we extend our basic model in two directions according to the suggestions of Smith and Malthus: we make agricultural productivity and the stock of the labour force L endogenous. In WN a stationary agricultural economy is characterized by a class of landowners that consume only luxury goods and a class of wage-earners that consume only food. Smith argues that a more egalitarian land distribution and different types of agricultural contracts may lead not only to a more egalitarian rent distribution among landowners but also to an increase of aggregate rent thanks to an increase of agricultural productivity (WN, III.ii). Both factors provide potential entrepreneurs with the incentives to invest in the manufacturing sector and may lead the economy out of long-run stagnation (see also Baldwin, 1956 and Strassman, 1956). Moreover, as we emphasized in section 5.2, Smith claims that the availability of a wider range of goods provides landowners with further incentives to increase the productivity of their lands.[9] To model the relationship between productivity in agriculture on the one hand and land distribution and availability of industrial goods on the other we suppose that a_F positively depends on equality in land distribution and on the range of industrial goods:

$$a_F = a_F(\lambda, Q), \frac{\partial a_F}{\partial \lambda} < 0, \frac{\partial a_F}{\partial Q} > 0 \tag{13}$$

where λ is an index of inequality (e.g. Gini index) of land distribution.

Our basic model neglects a crucial element of classical economics. We refer to the so-called Iron Law of wages and population (also called Malthusian Law of population) according to which a rate of wages above

subsistence ($w > z$ in our model) or below subsistence ($w < z$) leads to an increase (a decrease) in total population and agricultural employment. Formally:

$$\dot{L} = \Delta(w-z) \tag{14}$$

where $(d\Delta/dw-z) > 0$. Thus any $w-z$ gap is supposed to be a short-run phenomenon. *Rebus sic stantibus*, the economy described in our model would not be able to achieve long-run growth. Yet, as we have argued in Section 5.3, Malthus, whose name is usually strictly associated to the Iron Law, was perfectly aware that the availability of conveniences and luxuries may affect wage-earners' behaviour: given the size of their families, wage-earners may devote (at least partially) the excess of their income over subsistence to the consumption of conveniences. Thus, workers' population and workers' consumption of food may stay almost stationary or may grow not too fast in the face of increases in agricultural productivity and real wages. In terms of our model we must modify equation (14) and consider the change in labour force also as a function of the range of available conveniences:

$$\dot{L} = L(w-z, Q), \ \frac{\partial L}{\partial w-z} > 0, \ \frac{\partial L}{\partial Q} < 0 \tag{15}$$

In Sections 5.2 and 5.3 we briefly recalled Smith's and Malthus' analyses concerning the relations among income distribution, demand conditions and growth. In what follows we show that classical analysis may be easily formalized in terms of our model. In particular we discuss three cases which best fit our previous results. The first concerns the factors which prevent development in an agricultural economy. The second regards the transition from a stationary agricultural economy to an industrialized one and, finally, the third case concerns the conditions to be satisfied for an economy to achieve positive long-run growth.

Case I: a stationary agricultural economy

Consider an agricultural economy (i.e. an economy with no manufactures of the industrial type) characterized by a low level of agricultural productivity, that is, $w < F(L_F, a_F)/L_F < z$. In such an economy there is no demand for industrial goods if condition 4) does not hold, that is, if $\underline{s}(R+\Pi)+w < z$. In this case income is entirely devoted to food consumption and the conditions for take-off do not occur. This kind of poverty trap may happen if land property is too fragmented. A certain degree of inequality in land distribution may create the conditions for take-off provided that it generates sufficient expenditure out of rents for industrial goods (see equation (10)). An increase in agricultural productivity may lead to $w < z < F(L_F, a_F)/L_F$. In this case, the increase in agricultural productivity may lead to an increase in aggregate rent and create the physical resources to be potentially employed in manufactures. If inequality (4) holds, then manufactures may develop provided that land distribution is such that the economy generates a level of demand able to cover the fixed and variable costs of production of industrial goods. Thus if $N < N^*$, that is, if land property is too concentrated then demand for industrial goods is insufficient to make investment in manufactures profitable. In such a situation the increase in aggregate rent, generated by the increase in agricultural productivity, is spent on personal services and/or consumption of ‘finer’ foreign commodities. Domestic manufactures do not arise.

In this regard it may be objected that an increase in agricultural productivity may lead to an increase in wages and rents and not just rents. If an increase of $F(L_F, a_F)/L_F$ leads to $w > z$, then wage-earners may become industrial goods consumers provided that the so-called Iron Law of wages and population does not hold. As is well known, according to this law any increase in the rate of wages above subsistence (historically determined by ‘habits and customs’) eventually leads to an increase in workers’ population. In our model this means that w would exceed z only temporarily (see Case III below).

Case II: from a stationary agricultural economy to an industrial one

A more egalitarian land distribution leads to an increase in N and, therefore, a decrease in λ. An increase in N may lead to $N > N^*$ thus making investment in manufactures profitable. A decrease in λ positively affects a_F and thus leads to an increase in rents and wages (see equations (2) and (3) and the sign of the relevant derivatives). The above may be considered as the first phase of development. The second phase begins when manufactures start producing and thus industrial goods start being widely consumed in the economy. In this second phase the relation between a_F and Q becomes the crucial factor. The widening of the consumption set, that is, the increase in Q makes landowners willing to raise their disposable income. Landowners are thus ready to accept the introduction of more efficient agricultural techniques (even if this involves a partial loss of power over their tenants). Agricultural productivity increases and this leads to an increase in rents and wages. Finally, the increase in labour and property incomes leads to an increase in the number of industrial goods produced in the economy (see equation (10)). The growth of agricultural production and the growth of industrial production sustain each other: a virtuous circle of growth starts.

It is worth remarking that the development of a new productive sector requires a shift of the labour force from the traditional sectors of the

economy, that is, a shift from agriculture and personal services to manufactures. We have argued that manufactures start developing as soon as a decrease in λ causes an increase in a_F and, consequently, an increase in R and w. Given total population and the related total demand for food, an increase in a_F implies a labour force surplus in agriculture. Labour resources are thus set free to move to new developing manufactures.

Case III: long-run growth

In the previous subsection we showed how a stationary agricultural economy characterized by a low level of agricultural productivity may start its process of industrialization. Land redistribution and/or an increase in agricultural productivity may encourage shareholders to boost their demand for industrial goods. Provided that such a demand reaches a critical level, it becomes profitable for some agents within the economy to become entrepreneurs, that is, to invest resources in the refinement of domestic manufactures or in the domestic production of those ‘finer’ commodities previously imported from abroad. Shareholders’ demand thus plays a fundamental role in the take-off of an agricultural economy. Yet the explanation of long-run growth must include other factors. As highlighted by the modern literature on growth, technological progress is undoubtedly one of the crucial elements affecting long-run growth. While we leave the analysis of technological progress to future research, in what follows we focus on the level of wages and the related consumption choices of wage-earners. In Section 5.4 we assumed that shareholders constitute only a small fraction of total population. Thus, in the long-run, profits and employment in the manufacturing sector of the economy may be supposed to depend heavily on wage-earners’ expenditure on industrial goods. In terms of our model this means that in the long-run $(L - N)(w - z)$ becomes the crucial addendum in equations (11) and (12).

In our model an increase in a_F implies an increase in w (see equation (2) and the relevant derivative) and, possibly, an increase in w over z. Wage-earners start consuming industrial goods. But according to the Iron Law, a positive $w - z$ gap fosters a rise of total population, L. An increase in L implies an increase in food consumption. Accordingly, agricultural production and L_F rise. But an increase in L_F implies a decrease in w (see equation (2) and the relevant derivative) and thus a decrease in the positive $w - z$ gap. This fact would have negative effects on the growth performance of the economy since it is precisely $(L - N)(w - z)$ which mainly supports long-run growth if N is small in relation to L. According to equation (15) an increase in the variety of conveniences and luxuries, Q, may at least partially balance the growth of population, $\dot{L}$, driven by the excess of wage over subsistence, $w - z$. Total population may grow at a lower rate than that predicted by the Iron Law. Thus a positive $w - z$ gap may become persistent and display its positive effects on long-run growth.

5.7. FINAL REMARKS

It would not be an overstatement to say that contemporary growth literature lacks a systematic analysis of the relation between income distribution, consumption patterns, development and growth. This neglect of demand factors may prove a serious lacuna: the long-run stagnation of an agricultural economy as well as the long-run growth of an industrial economy may escape full comprehension. Our paper has tried to make good this lacuna. We have elaborated a simple model, basically an extension of Murphy et al. (1989), to examine: (*i*) some of the causes which force an agricultural economy into a situation of long-run stagnation; (*ii*) some of the basic forces driving an economy during its transition from the agricultural stage to the industrial one and, finally; (*iii*) some of the conditions to be fulfilled for an industrial economy to perform a long-run growth.

An agricultural economy may not escape from a poverty trap if agricultural productivity is too low. One of the possible causes of low agricultural productivity is excessive fragmentation of land property. Since in our model agents spend the whole of their income on food if income is lower than a certain threshold, neither wage-earners nor shareholders buy industrial goods. If agricultural productivity rises, rents and/or wages raise. The economy may thus be able to generate a level of expenditure for industrial goods which makes their production profitable. Yet an increase in agricultural productivity does not necessarily lead the economy out of its poverty trap. If only rents rise and land property is too concentrated, then extra-income may be spent on the consumption of personal services and/or foreign luxury goods. If only wages rise and the Iron Law of wages and population holds, then extra income is spent on extra consumption of food in order to rear a growing family. In both cases domestic manufactures do not develop.

However, if landowners acquire a strong taste for industrial goods they would strive to increase their disposable income. Thus they may be willing to implement a change in agricultural techniques which raise agricultural productivity. In such a scenario a virtuous circle of development may start: the growth of the agricultural sector, led by the growing availability of (and desire for) industrial goods, releases new resources for the growth of the manufacturing sector. Similarly, if wage-earners acquire a strong taste for industrial goods they would not strive to increase the share of their income devoted to food consumption. Thus they resist the temptation to increase the

size of their family and/or to reduce their labour supply. We have argued that if shareholders constitute only a small portion of total population, then the long-run growth performance of an economy depends on wage-earners' consumption patterns.

To conclude, we are well aware of several serious shortcomings in our work. To mention a few: the model presented here is static and its dynamic properties are still to be carefully studied. Moreover, we have not analyzed a most important factor in the long-run, technological progress. Finally, we have not considered capital accumulation. Much work is thus left over to future research.

NOTES

1. As remarked by Rosenberg: 'In a society where the finer manufactures are not available, opportunities for cultivating one's vanity are necessarily limited. In the absence of such commodities, large rental incomes are employed in hospitality, in the maintenance of a large group of retainers, and in acts of bounty to one's tenants. In spite of these acts of generosity, however, the typical behaviour of large landowners as late as the time of European feudalism was reasonably frugal. Large landowners were not extravagant, and it was even common for them to save' (Rosenberg, 1968, p. 367).
2. The distinction between the 'limiting principle of profits' and the 'regulating principle of profits' is drawn by Malthus in Book I, ch. V of PPE. See also Costabile and Rowthorn (1985).
3. As is well-known, Malthus' overall position on foreign trade is many-sided and not easy to grasp. We shall not deal with it in greater details since in our paper we consider a closed economy. Needless to say, the analysis of foreign trade could be an interesting extension of our model.
4. An example of agent's preferences compatible with the Engel's curves used in the text is the following:

$$U = \begin{cases} c & \text{for } c \le z \\ z + \int_0^\infty x(q)\, q^{-\gamma} dq & \text{for } c > z \end{cases}$$

   where $x(q) \in \{0,1\}$, $q \in [0,\infty)$, c is the food consumption, $z > 0$ is the threshold or the minimum amount of food consumption required before the consumption of non-agricultural goods begins; q is the index of the non-agricultural goods produced by the economy and $x(q)$ assumes value 1 if good q is consumed or 0 otherwise. Since marginal utility of good q is equal to $1/q$, agents prefer to consume goods with a lower index.
5. Notice that we do not assume that factors are paid their marginal productivity.
6. To clarify this point consider the following example: someone willing to hear a Mozart's symphony either may buy a CD or may buy a ticket to hear an orchestra playing live music.
7. The interpretation given in the text may be questionable, since generally the cost of innovation is payed once, while the return to investment is given by a flow of future profits. In our framework it is particularly difficult to model this aspect because future profits

depend on the extent of demand, while the latter depends on future innovations, changes in income distribution, etc. Zweimuller (2000) provides an analysis of this case, though limited to balanced growth equilibria.

8. To prove it consider that assumption (4) implies that $\underline{s}(R+\Pi)S^M > (z-w)S^M$, from which

$$\underline{s}(R+\Pi)S^M > (z-w)\left[\underline{s}\left(N-N^*\right)+\int_{\underline{s}}^{s^*}(s-\underline{s})\,dG(s)\right]$$

and finally:

$$(R+\Pi)S^M-\left(N-N^*\right)(w-z) > \frac{(z-w)\int_{\underline{s}}^{s^*}(s-\underline{s})\,dG(s)}{\underline{s}} > 0$$

9. Malthus closely follows Smith's argument: 'Adam Smith has well described the slack kind of cultivation which was likely to take place, and did in fact take place, among the great proprietors of the middle ages. But not only were they bad cultivators/ and improvers; and for a time perhaps deficient in a proper taste for manufactured products; yet, even if they had possessed these tastes in the degree found to prevail at present, their inconsiderable numbers would have prevented their demand from producing any important mass of such wealth. We hear of great splendour among princes and nobles in every period of history. The difficulty was not so much to inspire the rich with a love of finery, as to break down their immense properties, and to create a greater number of demanders in the middle ranks of life who were able and willing to purchase the results of productive labour. This, it is obvious, could only be effected very gradually. That the increasing love of finery assisted considerably in accomplishing this object is highly probable; but these tastes alone, unaccompanied by a better distribution of property in other respects, would have been quite inefficient' (Malthus, 1986, pp. 298–99). See also Maccabelli (1997).

6. Keynesian theories of growth

Pasquale Commendatore, Salvatore D'Acunto, Carlo Panico and Antonio Pinto

6.1. INTRODUCTION

This paper outlines the content of a Keynesian approach to the theory of growth. While for other established traditions it is possible to talk of a theory of growth described by some specified models and contributions,[1] for the Keynesian tradition it is only possible to identify several lines of development, which share the view that the economic system does not tend necessarily to full employment and that the different components of demand may affect the rate of growth of the economy.

As far as we know, there is no essay in the recent literature which seeks to reconstruct the content of a Keynesian approach to growth by describing the lines of research, which have historically emerged. In what follows an attempt will be made to do so. This attempt outlines a unified framework that can deal with the influence of the different components of aggregate demand on the rate of growth of an economic system that does not tend necessarily to full employment. The specification of this unified framework makes it possible to preserve the diversity of the ideas proposed by Keynesian authors on what can be considered the most relevant factors at work.[2] Moreover, it shows Keynesian growth theorists as a homogeneous crew, sharing a positive theoretical standpoint on the role of aggregate demand, rather than a group of authors united by a critical attitude towards orthodoxy, but unable to present a systematic challenge to the dominant theories.[3]

The paper is so organised. Section 6.2 aims to derive a unifying framework for Keynesian theories of growth from the analyses proposed by Harrod, the founder of modern growth theory. Sections 6.3, 6.4 and 6.5 deal with the analyses underlining the influence on growth of three components of effective demand, coming from the Government sector, the private sector, in the form of autonomous investment (i.e. investments not directly generated by savings), and the foreign sector. Section 6.6 draws some conclusions.

6.2. HARROD AND THE FORMATION OF A KEYNESIAN FRAMEWORK FOR GROWTH THEORY

According to Varri (1990, p. 9), Harrod's contributions to growth have received less attention than they deserve. Recently, however, Young (1989) and Besomi (1999) have reconsidered his writings, taking advantage of the availability of his papers at the Chiba University of Commerce in Ichikawa (Japan) and clarifying the extent to which some of his writings have been misrepresented. They have refuted, in particular, the view that Harrod's efforts to develop a theory of growth and dynamics were stimulated by his work on imperfect competition and his dissatisfaction with the Austrian trade cycle theory put forward by Hayek (see Kregel, 1980, p. 98; 1985, pp. 66–7). Moreover, they have confirmed the limits of the widespread belief that Harrod developed his analysis of growth by assuming absence of monetary influences and fixed technical coefficients and saving propensity, in order to establish the famous 'knife-edge problem' (Solow, 1956, 1970; for the opposite interpretation, see Eisner, 1958, Asimakopulos and Weldon, 1965, Kregel, 1980, Asimakopulos,1985).

In opposition to the first view, Young (1989, pp. 15–50) clarified that Harrod's efforts to develop a theory of growth and dynamics were mainly stimulated by his contacts with Keynes. These began in 1922, when Keynes invited Harrod to study economics in Cambridge under his supervision (Phelps Brown, 1980, pp. 7–8). One year later, having read *A Tract on Monetary Reform*,

> Harrod took up Keynes's call for deeper research into the problems of the 'credit cycle', and over the next few years produced a number of essays on the subject. In these Harrod focused on the theoretical basis for – and policy options related to – issues raised by Keynes in the *Tract* (Young, 1989, p. 16).

According to Young, in these essays, some of which were never published, Harrod dealt with a problem that was central to Keynes's and other works of the time. Moving on from the idea that the economic system is stable and that negative influences on fluctuations only come from monetary and credit factors, attempts were made to identify a 'neutral' policy, i.e. a policy that can prevent monetary and credit disturbances from amplifying the fluctuations of the economy.

In those years Harrod also focused on Keynes's proposals for Government interventions.[4] According to Phelps Brown (1980, pp. 13 and 18), Harrod first heard Keynes's proposals at the Liberal Summer School of August 1924.[5] From then onwards, he closely followed Keynes's intellectual

activity on this subject and after the Great Depression he actively supported Keynes's proposals.[6] By that time, Harrod had come to recognise the need for deep political and theoretical changes. As Young (1989, pp. 30–8) points out in an unpublished paper written in 1933, Harrod stated that the Great Depression had posed a new problem to economists and politicians. The previous recessions had not led the economy too far from full employment, nor had they cast doubts on the belief that the economy is able to return to it. The severity of the Great Depression had changed this situation. It had jeopardised political stability and raised the problem both of a new political approach and of a new economic theory able to clarify whether market forces can lead the economy towards full employment or Government intervention is required to restore it.

As an initial contribution to these problems in 1933 Harrod published *International Economics*. This book, as Young (1989, pp. 38–9) points out, sets the lines of analysis that Harrod developed in the following years. In *International Economics* and in his 1936 *The Trade Cycle*, he moved from Keynes's *Treatise* (Young, 1989, pp. 48–50), to focus on the cyclical fluctuations of the economy around a line of steady growth. His aim was to point out that competitive market forces may widen the gap between actual and equilibrium growth, independently of the destabilising influences of monetary and credit factors, which had been underlined by the literature of the time. His 1939 essay on dynamics, again stimulated by the discussions with Keynes (*CW* XIV, pp. 150–79), focused instead on the equilibrium paths of the economy and on the factors determining the 'warranted' and the 'natural' rates of growth. This study represented 'a preliminary attempt to give an outline of a "dynamic" theory' (Harrod, 1939, p. 254) and 'a necessary propaedeutic to trade–cycle study' (p. 263).

It moved from the condition of equilibrium in the commodities' market. In the most simplified case, that of an economic system without Government intervention and closed to non-residents, this condition is represented by the equality between saving and investment decisions. In the formal presentation of his analysis, the saving propensity was taken as given. Yet Harrod (1939, p. 276) made some reference to the influence of the interest rate on the propensity to save and, in his following writings, he recalled the possibility of using Ramsey's intertemporal approach on which to base this part of his analysis.[7] The equation relative to investment, which introduces, according to Sen (1970, pp. 11 and 23) and Asimakopulos and Weldon (1965, p. 67), the major difference with other traditions, assumes that investment decisions are taken independently of saving decisions and are not generated by them. They depend on the 'acceleration principle' and on the degree of utilisation of capital equipment, along the following lines:

$$i = k\,g^* + f(g - g^*_{-1}) \tag{1}$$

where $f(0) = 0$ and $df/dg > 0$, i is the ratio between investment and the net output of the economy, g^* is the current period expected rate of growth of output, g^*_{-1} is the previous period expected rate of growth, g *is* the current period rate of growth, k is the equilibrium capital/output ratio.

Harrod used his analysis to study the 'warranted' rate of growth (g_w), defined as that equilibrium rate which allows the normal utilisation of capital equipment.[8] He assumed that, along the warranted equilibrium path, expectations are realised $(g^*_{-1} = g)$ and the expected rates of growth are equal to the warranted rate $(g^* = g^*_{-1} = g_w)$. The following equations were thus used for the analysis of the warranted rate:

$$s = k\,g_w \tag{2}$$

$$k = k(r), \quad (k'(r) \leq 0) \tag{3}$$

$$r = r_0 \tag{4}$$

where s is the average propensity to save and r is the rate of interest.
The introduction of equation (3) and (4) points out, in opposition to a widespread view, that Harrod did not develop his analysis of growth by assuming absence of monetary influences and fixed technical coefficients. Equation (4) assumes that the rate of interest depends on the conduct of monetary policy, which, according to Harrod, operates by stabilising this rate at some specified level.[9] Equation (3) recognises the possibility of substitution between factors of production. Harrod admitted the existence of decreasing marginal returns,[10] but considered that this kind of substitution was low, following the results reached by the Oxford Research Group, in which he actively participated.

From equation (2) one can derive

$$g_w = \frac{s}{k} \tag{5}$$

The study of the 'warranted' rate was for Harrod a preliminary part of the analysis of the dynamic behaviour of the economy, which in 1939 was presented through the following steps.

The first step dealt with the forces that start to operate as soon as the economy gets out of equilibrium and expectations are not realised. According to Harrod (1939, pp. 263–7), when the rate of growth differs from the equilibrium warranted rate, some centrifugal forces operate. If the former exceeds the latter, capital equipment is utilised above its normal level, inducing entrepreneurs to increase their investment decisions, as pointed out by equation (1). In the opposite case, capital equipment is utilised below its

normal level, inducing entrepreneurs to reduce investment decisions. In both situations, the rate of growth will be pushed further away from the warranted level. This description was considered by Harrod (1939, pp. 263–4) equivalent to that developed by static theory when it is assumed that the market price exceeds (is lower than) the equilibrium price and the appearance in that market of an excess supply (an excess demand) tends to restore equilibrium. These descriptions, unlike the 'cobweb' analysis in the traditional supply and demand theory, do not represent a dynamic analysis of disequilibrium. They just point out in an informal way that some centrifugal or centripetal forces come into operation as soon as disequilibrium occurs.

Most literature has interpreted this part of Harrod's work as the outcome of a dynamic analysis of stability. Sen (1979, p. 14), for instance, after pointing out that Harrod's analysis only deals with the initial elements of this problem and can be compatible with different analytical developments, criticised his conclusions.

> There are many other ways in which Harrod's somewhat incomplete model can be completed. Some confirm instability, while others either eliminate it or make it conditional on certain actual circumstances. In general, it will be fair to say that Harrod's instability analysis over-stresses a local problem near the equilibrium without carrying the story far enough, and extensions of his model with realistic assumptions about the other factors involved tend to soften the blow (Sen, 1970, p. 14).

Already in 1939, however, Harrod had stated that his analysis did not give a complete account of the problem, suggesting some lines along which a dynamic analysis of the behaviour of the system can be developed.

> Space forbids an application of this method of analysis to the successive phases of the trade cycle. In the course of it the values expressed by the symbols on the right-hand side of the equation undergo considerable change. As the actual growth departs upwards or downwards from the warranted level, the warranted rate itself moves and may chase the actual rate in either direction. The maximum rates of advance or recession may be expected to occur at the moment when the chase is successful (Harrod, 1939, pp. 271–2).

Moreover, in the subsequent years, Harrod (1948, p. 99) first claimed that he was reluctant to enter the field of the dynamic analysis of disequilibrium without developing the analysis of the equilibrium warranted path which, according to him, had a higher degree of generality.[11] He then rejected the view that his aim had been to raise a 'knife-edge problem'[12] confirming that he had only tried to underline the existence of some centrifugal forces coming into play as soon as the economy gets out of equilibrium. The reference to these forces did not exclude the existence of other forces, producing stabilising effects, which have to be analysed by considering, according to Harrod, that the ‘natural’ rate of growth represents the ‘ceiling’ limiting the expansion of the economy.

The second step of the analysis proposed by Harrod (1939) to study the dynamic behaviour of the economy considered the existence of forces pushing the ‘warranted’ rate of growth towards the ‘natural’ rate. This part of Harrod‘s work was based on his assumptions on substitution between factors of production and on the determination of the interest rate. As stated above, Harrod did not deny the existence of substitution between factors of production, but considered that it occurred to a small extent. After 1939, this idea was often restated: he claimed, with increasing emphasis, that he was skeptical on the possibility of reaching full employment through reduction of the interest rate.[13] Moreover, he confirmed that the rate of interest tends to show some rigidity, since it depends on the conduct of monetary policy, which, according to Harrod (1948, pp. 99–100; 1973, p. 67), operates by stabilising this rate at some specified level. This view of the interest rate, which also took into account the attempts of the monetary authorities to maintain the equilibrium of the balance of payments (Harrod, 1969, pp. 178 and 191; 1973, p. 75), raises the problem of the links between the theory of growth and that of distribution, since it was associated in Harrod‘s writings with the idea that a persistent change in this rate leads to a similar variation in the rate of profit.[14] The analysis of this problem, however, was little developed by the Oxford economist, who focused instead on the conclusion that one cannot rely on the belief that the spontaneous operation of market forces always leads the economic system towards full employment.

This conclusion led to the third step of analysis relative to the role of effective demand and Government policy on growth. Harrod (1939) pointed out that the warranted rate could be influenced by three different components of effective demand coming from the Government sector, the private sector, in the form of autonomous investment, and the foreign sector. Harrod (1939, pp. 269–74) gave some initial formal account of how these three sources of demand can affect the equilibrium path of the economy. Then, he focused on the Government sector and considered how policy can be used to stabilise the economy and to achieve higher growth and employment.

To sum up, the recent studies on Harrod‘s papers clarify that his seminal work on growth theory and dynamics was conceived as an extension of Keynes‘ analysis to a long-period context. It developed the view that the economic system does not tend necessarily to full employment and that the different components of aggregate demand may affect the rate of growth of the economy. His theory can be considered a prototype of a Keynesian

approach to this problem: it outlines a framework that much literature within this tradition has subsequently adopted.

6.3. THE INFLUENCE OF THE GOVERNMENT COMPONENT OF AGGREGATE DEMAND

The need to take into account the influence of Government activity on growth was pointed out by Harrod (1939, pp. 269–70 and 275), who also gave some initial formal account of how this source of demand can affect the equilibrium growth path of the economy. For him, Government policies have to be used both to stabilise the economy and to achieve higher growth.

> Policy in this field is usually appraised by reference to its power to combat tendencies to oscillations. Our demonstration of the inherent instability of the dynamic equilibrium confirms the importance of this. But ... in addition to dealing with the tendency to oscillation when it occurs, it may be desirable to have a long-range policy designed to influence the relation between the proper warranted rate of growth and the natural rate (Harrod, 1939, p. 275).

In 1939 Harrod claimed that both fiscal policy and variations in the long-term interest rate have to be used to pursue this long-range objective, adding that the latter are more appropriate than the former to this aim. The bank rate policy can be used instead to combat the runaway forces of the economy.

> If permanent public works activity and a low long-term rate availed to bring the proper warranted rate into line with the natural rate, variations in the short-term rate of interest might come into their own again as an ancillary method of dealing with oscillations (Harrod, 1939, p. 276).[15]

This position was maintained in Harrod (1948, pp. 74–5 and 117–22), where he again identified fiscal policy with ‘public works’. In the subsequent writings these ideas were revised, claiming that it was advisable to rely on fiscal, rather than on monetary policy, to affect the equilibrium warranted path, so as to bring it close to the natural path, and to conduct fiscal policy by changing the tax rates while keeping Government expenditure constant.

This new position was presented in Harrod (1964 and 1973), where he also recalled that the conduct of policy is difficult owing to the complexity of the objectives to be achieved (Harrod, 1964, pp. 913–15) and to the fact that

> even if the authorities had succeeded in maintaining a steady growth rate ... for a substantial period of time – a state of affairs not yet realised – and there was general confidence that their success would continue, this would not relieve the entrepreneur of his major uncertainties ... Entrepreneurs usually have to cast their bread upon the water (Harrod, 1964, p. 907).

He proposed to use the equilibrium condition of the commodity market to study how Government policy has to be applied and suggested dealing with this equation by taking the natural rate of growth as given, i.e. as the objective that the long-term policy has to pursue. Harrod (1973, p. 45) considered Government intervention necessary, arguing that this view was becoming increasingly popular.

> In the spectrum of countries ranging from individualism to socialism, the U.S.A. may be regarded as being at or near the individualist end. But even in that country 'monetary' and 'fiscal' policies are regarded as legitimate weapons of government, including the central bank. These policies serve to doctor the saving ratio and to provide enough, neither more nor less, to maintain reasonably full employment and growth in accordance with the growth potential of the economy (Harrod, 1973, pp. 28–9; see also 1964, p. 906).

He also underlined that the traditional position, which confines the use of these policies only 'to ironing out the business cycle', 'implies too narrow a view of the duties of the authorities' (Harrod, 1973, p. 29).

Finally, Harrod (1964, p. 906; 1973, pp. 102–3, 173 and 177) claimed that fiscal policy was appropriate to achieve this long-term objective. It should be used by varying the tax rates while keeping government expenditure constant (Harrod, 1973, p. 107). Monetary policy was appropriate instead to deal with what he defined the short-term policy objective of correcting the divergence of the actual rate from the warranted rate and stabilising the fluctuations of the economy. Temporary variations in the short-term rate of interest operate through their effects on the availability of credit in the markets (i.e. credit rationing) (Harrod, 1964, pp. 912–3; 1973, pp. 178–9). On the other hand, permanent variations in the interest rate tend to be more effective in causing similar variations in the rate of profit than in changing the capital–output ratio (Harrod, 1973, pp. 44, 78 and 111).

The formal analysis used by Harrod to deal with these views was limited. It can be developed as done in equation (6) below, which follows his proposal to study how to apply Government policy by using the equilibrium condition of the commodities' market, which in this case takes the form 'saving plus taxation is equal to investment plus Government expenditures'.

$$s\,(1 - t + r_b b) + t = kg + h + r_b b \tag{6}$$

where s is the private sector's propensity to save ($0 < s < 1$), t is the average tax rate, defined in terms of the net output of the economy ($0 < t < 1$), r_b is the interest rate on Government bonds, b is the amount of Government bonds in circulation, measured in terms of the net output of the economy ($b \geq 0$), k is the capital–output ratio ($k > 0$), g is the rate of growth of the

economy, h is the amount of Government's expenditure on goods and services, measured in terms of the net output of the economy ($h \geq 0$).

As Harrod suggests, this equation can be used either to study the factors affecting the warranted rate of growth (in this case, g is taken as unknown, while r and the policy parameters t and h are taken as given) or to analyse how fiscal policy has to be applied to maintain reasonable full employment or growth in accordance with the potential of the economy (in this case, g is taken as given at its natural level, while one policy parameter, say t, is considered unknown).

From equation (6) one can derive

$$g = \frac{s(1 - t + r_b b) + t - h - r_b b}{k}\,. \tag{7}$$

It can be noticed that variations in the tax rate keep affecting growth even in the simplified case of a balanced Government budget and absence of Government bonds ($t = h > 0$ and $b = 0$), when equation (7) becomes

$$g = \frac{s(1 - t)}{k}\,. \tag{8}$$

The influence of t on g does not depend on that of t on the propensity to save and on the capital–output ratio [16]

The presence of Government debt and the interest rate in equation (7) raises the problem of the relationships between growth and distribution and between monetary and fiscal policy. Only the former problem is known to occupy a central place in the original development of the post Keynesian theory of growth and distribution.[17] Kaldor's 1958 Memorandum to the Radcliffe Committee, however, considers both problems simultaneously.

The Memorandum describes how Government policy can affect stability and growth. It argues that monetary policy has to stabilise the short-term interest rates in order to avoid some 'undesirable consequences'. The instability of the interest rates enhances financial speculation and reduces the ability of the markets to convey financial resources towards productive enterprises. Moreover, it raises the risk premium to be paid on loans of longer maturity and leads to higher long-term interest rates. Higher long-term interest rates, in turn, make the management of Government debt difficult. Moreover, they increase the probability that firms may not be able to pay back their loans, making lending institutions and financial markets more fragile. Finally, they tend to cause economic stagnation.

To justify the tendency to stagnation Kaldor made reference to his theory of growth and distribution and to the 'Cambridge equation'.

> In a steadily growing economy the average rate of profit on investment can, in the first approximation, be taken as being equal to the rate of growth in the money value of the gross national product divided by the proportion of profit saved … To keep the process of investment going, the rate of profit must exceed the (long-term) interest rates by some considerable margin (Kaldor, 1958, pp. 137–8)

A monetary policy causing unstable interest rates raises the long-term rates to a level considered by investors too high to keep accumulation going. Under these circumstances, stagnation prevails, unless the rate of profit is raised too. According to Kaldor, this can be done through fiscal policy.

> If the rate of interest were higher than [the level that keeps investment going], the process of accumulation would be interrupted, and the economy would relapse into a slump. To get it out of the slump it would be necessary to stimulate the propensity to consume – by tax cuts, for example – which would raise the rate of profit and thus restore the incentive to invest (Kaldor, 1958, p. 138).[18]

The post Keynesian theory of growth and distribution, to which Kaldor greatly contributed, differs from Harrod‘s growth theory for the introduction of the saving propensities of different income groups and for the role attributed to distributive shares in restoring equilibrium conditions. According to some literature, this part of Kaldor‘s work departs from the Keynesian tradition, since it does not reject the idea that market economies tend to full employment.

Kaldor‘s Memorandum to the Radcliffe Commission does not confirm this allegation (Kaldor, 1958, pp. 135–7 and pp. 141–2). It shows many similarities with the views proposed by Harrod and the rest of Keynesian tradition on the role of Government policy. First of all, Kaldor considered Government policies necessary to pursue stability and growth. Secondly, thought that Government policies have to deal with a complex set of objectives, which are interrelated – and often incompatible – among them. Thirdly, for Kaldor, monetary policy is the appropriate tool against the fluctuations of the economy, while it is advisable to use fiscal policy to pursue the long-range objective of sustained growth. Fourthly, when he advocated fiscal policy, Kaldor referred to variations in the tax rate, rather than to variations in the level of Government expenditure. Finally, like Harrod, Kaldor proposed to use the equilibrium condition of the commodities’ market to deal with these problems and referred to it either to determine the growth path of the economy (considering the rate of growth as unknown and the interest rate, the tax rate and Government expenditure as given) or to determine the intensity of fiscal policy appropriate to the achievement of a specific rate of growth (considering one policy parameter – the tax rate – as unknown and the rate of growth as given).

Kaldor did not present his positions on the role of Government policy in a formalised way. Nor can such a treatment be found in other literature of that time. His reference to the Cambridge equation must then be considered, as he himself stated, a first approximation rather than the result of a thorough treatment of this problem. The first formal presentation of the post Keynesian theory of growth and distribution, which explicitly introduced the Government sector, was provided by Steedman (1972). This article proved that in an analysis that assumes a balanced Government budget and no outstanding bonds, the Cambridge equation holds in a larger number of cases than the ‘dual theorem’ of Modigliani and Samuelson. Some years later, Fleck and Domenghino (1987), who challenged the validity of the Cambridge equation when the Government budget is not balanced, stimulated an intense debate on this subject. The debate has examined a large number of cases, showing when the Cambridge equation holds and confirming the conclusion that Steedman had previously reached.[19]

The results of the debate show how the views on the role of Government policy that Kaldor presented in the Memorandum to the Radcliffe Commission can be formally developed and clarify some features of his proposals. Let us consider the case examined by Denicolò and Matteuzzi (1990), in which the Cambridge equation holds. It refers to a closed economy with two classes (workers and capitalists),[20] where the Government sector finances its budget through the issue of bonds and the private sector finances its productive activity through the sale of shares to other components of the private sector. Capitalists do not work: they earn their income through the returns of their wealth. Moreover, the two classes have different saving propensities, can invest their wealth in shares representing real capital and in Government bonds, and have the same portfolio structure (for the case of different portfolio structures, see Panico, 1993). To study what are the conditions allowing steady growth, we must specify the equilibrium condition in the commodities’ market, the dynamic equilibrium conditions between the savings of the two classes and the growth of their wealth, and the dynamic equilibrium condition between the Government budget and its debt. These conditions can be written as follows:

$$s_c(1-t)\,\alpha(r_b b + r_k k) + s(1-t)\,[1 + r_b b - \alpha(r_b b + r_k k)] + t = gk + h + r_b b \quad (9)$$

$$s_c(1-t)\,\alpha(r_b b + r_k k) = g\,\alpha(b + k) \quad (10)$$

$$g\,b = h + r_b b - t \quad (11)$$

where s_c is the propensity to save of the capitalist class ($0 < s_c < 1$), t is the tax rate ($0 < t < 1$), which is assumed to be the same on all forms of income, α is the quota of wealth owned by the capitalist class ($0 \leq \alpha \leq 1$), s_w is the propensity to save of the working class ($0 < s_w < s_c$), r_b is the rate of interest on bonds, b is the stock of Government bonds measured in terms of the net output of the economy ($b \geq 0$), g is the rate of growth, k is the capital/output ratio ($k > 0$), h is the Government expenditure on goods and services, measured in terms of net output ($h \geq 0$), r_k is the rate of return on real capital. If we assume $r_b = r_k = r$, equation (10) becomes:

$$s_c(1 - t)r = g. \tag{12}$$

This confirms the validity of the Cambridge equation, taking into account the role of t, and allows one to calculate the value of t compatible with steady growth at the rate of interest fixed by the monetary authorities.

Equations (9)–(12) thus show how to develop in a formal way the views proposed by Kaldor in his Memorandum to the Radcliffe Commission, where the lack of a formal analysis of how Government intervention can affect growth and distribution led the author to refer to a version of the Cambridge equation which, unlike equation (12), does not include the tax rate. As a consequence, Kaldor conceived the influence of tax variations on growth in terms of their effect on the propensities to save. The analysis presented above, instead, clarifies how Government intervention can affect demand and growth independently of changes in the propensities to save and in the capital–output ratio. It thus further elaborates Kaldor‘s attempt to describe how fiscal policy can be used to maintain steady growth conditions.

Finally, the results of the recent debate on the role of the Government sector in the post Keynesian theory of growth and distribution clarify some other common elements of the classical and the Keynesian traditions (see Panico, 1997, 1999). They allow reconciliation of two approaches to distribution, which have been considered alternative (see Moss, 1978, p. 306; Vianello, 1986, p. 86; Nell, 1988; Pasinetti, 1988; Pivetti, 1988; Wray, 1988; Abraham–Frois, 1991, pp. 197 and 202). These are the approach proposed by Kaldor and Pasinetti in their theory of growth and distribution and that implied by Sraffa‘s suggestion in *Production of Commodities* to take the rate of profit, rather than the wage rate, as the independent variable in the classical theory of prices and distribution.

6.4. THE INFLUENCE OF AUTONOMOUS INVESTMENT

The introduction of an autonomous investment function is often considered what differentiates a Keynesian theory of growth from other approaches. There is, however, no agreement in the literature on what characterises a

Keynesian investment function and several investment-led growth theories have been proposed. The first type of theory (labelled neo-Keynesian) was proposed by Joan Robinson (1956, 1962) and Kaldor (1957 and 1961). They are characterised by full capacity utilisation of plants, flexible income shares and a functional relationship between the rate of capital accumulation and the rate of profits.[21] A second group of theories (labelled Kaleckian) was inspired by the works of Kalecki (1971) and Steindl (1952). They assume that firms under-utilise their productive capacity and apply mark-up procedures in determining prices. Moreover, capital accumulation is driven by profitability (through the rate of profits) and by effective demand (through the degree of capital utilisation). These investment-led growth theories have been further elaborated in the literature. In what follows, an attempt is made to compare the alternative lines of development of investment-led growth within the Keynesian tradition by introducing a homogeneous set of equations which can be modified to take account of the assumptions relating to capital utilisation, income distribution and investment determinants.

Let’s assume (*i*) a closed economy with no government intervention; (*ii*) two factors of production, labour and capital, with a fixed coefficient technology; (*iii*) flexible labour supply; (*iv*) absence of technological progress and capital depreciation; (*v*) identical physical composition of capital and product; (*vi*) homogeneous firms. The following equations can then be written

$$1 = wa_l + r_k k \tag{13}$$

$$\frac{1}{k} = \min\left(\frac{l}{a_l}, \frac{1}{a_k}\right) \tag{14}$$

$$u = \frac{a_k}{k} \tag{15}$$

$$\min(w_\pi, w_\omega) \leq w \leq \max(w_\pi, w_\omega) \tag{16}$$

$$s = s_c r_k k \tag{17}$$

$$\frac{i}{k} = \gamma(r_k, u, g) \tag{18}$$

$$s = i \tag{19}$$

where k is the capital/output ratio, w is the real wage rate, r_k is the rate of profits, l is the labour/capital ratio, a_l is the labour coefficient of production, a_k is the capital coefficient of production, u is the degree of capacity utilisation, w_π is the wage firms are prepared to pay, w_ω is the wage workers are prepared to accept, s is the ratio between saving and output, i is the ratio between investment and output, g is the rate of growth of income, s_c is the capitalists' propensity to save, with $0 < s_c \leq 1$.

According to equation (13) output (normalised to one) is distributed between wage and profit recipients. Following expression (14), which describes a fixed-coefficient (Leontief) type technology, the elastic labour supply guarantees that the labour/output ratio always coincides with the corresponding technical coefficient, $a_l = lk$. Conversely, capital is not necessarily fully utilised. It follows that output is not necessarily the maximum technologically possible, $1/k \leq 1/a_k$. Expression (14) leaves open the determination of the degree of capacity utilisation, defined in expression (15) as the ratio between current demand and full capacity output. It is possible to envisage two cases. In the first, capacity is fully utilised, that is, the equality $u = 1$ $(1/k = 1/a_k)$ holds. In the second, some capacity is left idle with the degree of capacity utilisation settling in any period at some level which does not necessarily equal one, that is, $u \leq 1$ $(1/k \leq 1/a_k)$. Expression (16) also leaves the wage rate open to two possible determinations. In the first case, workers' and firms' claims over the shares of income (in real terms) are not inconsistent, $w_\omega \leq w \leq w_\pi$. If follows that distribution and growth are simultaneously determined. In the second case, workers and firms lay conflicting claims over income shares, $w_\pi \leq w \leq w_\omega$ (and $w_\omega \neq w_\pi$). The distribution between profits and wages depends on the relative power of workers and firms. The way in which distribution is in fact determined depends on the institutional setting. Equation (17) clarifies that saving propensities differ between classes. According to expression (18), investment demand depends on profitability (through r_k), on the demand level (through u) and on demand growth (through g). Keynesian approaches to investment-led growth differ inasmuch as they do not assign to each of the determinants of investment the same prominence. Finally, equation (19) represents the equilibrium condition saving equal to investment. The model (13)–(19) has three degrees of freedom. The way in which it is closed differentiates the Keynesian approaches to investment-led growth.

The neo-Keynesian position is represented by the following equations derived from expressions (13)–(19) by assuming full capacity utilisation, $u = 1 (k = a_k)$; endogenous income distribution, $w_\omega \leq w \leq w_\pi$; and disregarding the role of the rate of growth of demand in the investment function:

$$1 = wa_l + r_k a_k \tag{20}$$

$$s = s_c r a_k \tag{21}$$

$$\frac{i}{k} = \gamma_0 + \gamma_1 r_k \tag{22}$$

$$s = i \tag{23}$$

By rearranging (20), one obtains the following expression

$$r_k = \frac{1}{a_k} - w\frac{a_l}{a_k} \tag{24}$$

which describes the traditional long-term negative relationship between r and w. Following Joan Robinson (1962), investors' 'animal spirits' (encapsulated in the constant coefficients γ_0 and γ_1) are prompted by expected profitability and favoured by the availability of internal finance. This explains the relationship (22) between desired investment and the rate of profits.

The model (20)–(23) is similar to that proposed by Marglin (1984a, 1985b) to describe the contributions of Joan Robinson and Kaldor to growth theory. By imposing the equilibrium growth condition according to which all the variables have to grow at the same rate, $i/k = g$, the solutions are univocally determined:[22]

$$\bar{r}_k = \frac{\gamma_0}{s_c - \gamma_1} \tag{25}$$

$$\bar{g} = s_c \frac{\gamma_0}{s_c - \gamma_1} \tag{26}$$

There are three major features of the neo-Keynesian analysis. The first is that distribution and growth are simultaneously determined. The second is the transposition to the long run of the so-called 'paradox of thrift', according to which an increase in the propensity to save induces a reduction in the rate of growth and in the equilibrium rate of profits. Indeed, by differentiating expressions (25) and (26) with respect to s_c one obtains

$$\frac{d\bar{r}_k}{ds_c} = -\frac{\gamma_0}{(s_c - \gamma_1)^2} < 0 \tag{27}$$

$$\frac{d\bar{g}}{ds_c} = -\frac{\gamma_0 \gamma_1}{(s_c - \gamma_1)^2} < 0 \tag{28}$$

The third is the negative relationship between g and w. From (21), (23) and (24), taking into account the equilibrium condition $i/k = g$, it follows that

$$\frac{d\bar{g}}{d\bar{w}} = -\frac{s_c a_l}{a_k} < 0 \tag{29}$$

Lower levels of the wage rate correspond to higher accumulation. Profit leads growth.

If the equilibrium solution $\bar{w}$ lies outside the interval $w_\omega \le w \le w_\pi$, the neo-Keynesian analysis becomes overdetermined. When the left constraint is binding, $w = w_\omega > \bar{w}$ and $r_k < \bar{r}_k$, the economy suffers inflationary pressures, because investment demand permanently exceeds saving, $\bar{g}a_k > i > s$. Joan Robinson (1962) acknowledged this possibility by referring to an ‘inflationary barrier’ (also named ‘real wage resistance‘), which represents the minimum level of the real wage rate organised labour is prepared to accept without opposing rises in monetary wages.[23] Conversely, when the right constraint is binding, $w = w_\pi < \bar{w}$ and $r_k > \bar{r}_k$, the economy is stagnating since investment is too low (or saving is too high) for full capacity growth, $s > i > \bar{g}a_k$. This constraint may become operational when, following Kaldor (1957a), firms are – regardless of demand – not prepared to lower prices below that level which guarantees a minimum profit margin π, which determines $w_\pi = (1/a_l) - \pi/a_l$ and depends on the Kaleckian ‘degree of monopoly’. Note that the discrepancy between s and i can be reduced by varying s_c or (in the opposite direction) γ_0 and γ_1.

Unlike the neo-Keynesian approach, some economists (e.g. Rowthorn, 1981; Dutt, 1984, 1987, 1990; Nell, 1985; Amadeo, 1986a, 1986b, 1987 and Lavoie, 1992, 1995), inspired by the works of Kalecki and Steindl, developed analyses in which firms are allowed to operate under long-run under-utilisation of production plants. In Kaleckian analyses demand affects capital accumulation through changes in the degree of capacity utilisation. They assume, moreover, oligopolistic markets and conflicting claims over income distribution, $w_\omega > w_\pi$. This position can be represented by the following equations derived from expressions (13)–(19) by assuming an endogenous degree of capacity utilisation, $u \le 1$; exogenous income distribution, $w = w_\pi$; and disregarding the role of the rate of growth of demand in the investment function:

$$1 = wa_l + r_k k \tag{30}$$

$$u = \frac{a_k}{k} \tag{31}$$

$$w = w_\pi \tag{32}$$

$$s = s_c r_k k \tag{33}$$

$$\frac{i}{k} = \gamma_0 + \gamma_1 r_k + \gamma_2 u \tag{34}$$

$$s = i \tag{35}$$

According to expression (32), income distribution is determined outside the model according to the Kaleckian theory of distribution. It is assumed that firms, independently of workers’ wage resistance, fix prices through a mark-up procedure securing profit margin π, wage rate $w_\pi = (1/a_l) - \pi/a_l$ and profit share $r_k k = 1 - w_\pi a_l = \pi$.[24] Moreover, using (31), a relationship may be expressed between the rate of profits and the degree of capacity utilisation,

$$r_k = \frac{\pi u}{a_k} \tag{36}$$

according to which r_k is not univocally determined by income distribution as it was, according to expression (24), in the neo-Keynesian model. Equation (34), a linear form of (18), postulates a relationship between capital accumulation, the rate of profits and the degree of capital utilisation, specified by the constant coefficients γ_0, γ_1 and γ_2.[25] In Kaleckian writings the current rate of profits is relevant for investment decisions for two main reasons. It represents a proxy for expected profitability and also a source of internal financing.[26] The level of capacity utilisation affects investment decisions both indirectly (acting through the rate of profits) and directly by reflecting the state of demand.[27]

By imposing the equilibrium growth condition $i/k = g$, the solutions of equations (30)–(35) are univocally determined:

$$\bar{u} = \frac{\gamma_0 a_k}{(1 - w_\pi a_l)(s_c - \gamma_1) - \gamma_2 a_k} \tag{37}$$

$$\bar{r}_k = \frac{\gamma_0 (1 - w_\pi a_l)}{(1 - w_\pi a_l)(s_c - \gamma_1) - \gamma_2 a_k} \tag{38}$$

$$\bar{g} = \frac{s_c \gamma_0 (1 - w_\pi a_l)}{(1 - w_\pi a_l)(s_c - \gamma_1) - \gamma_2 a_k} \tag{39}$$

Note that the paradox of thrift is preserved, as shown by differentiating expressions (38) and (39) with respect to s_c,

$$\frac{d\bar{r}_k}{ds_c} = -\frac{\gamma_0 (1 - w_\pi)^2}{[(1 - w_\pi a_l)(s_c - \gamma_1) - \gamma_2 a_k]^2} < 0 \tag{40}$$

$$\frac{d\overline{g}}{ds_c} = -\frac{\gamma_0(1-w_\pi a_l)[\gamma_1(1-w_\pi a_l)+\gamma_2 a_k]}{[(1-w_\pi a_l)(s_c-\gamma_1)-\gamma_2 a_k]^2} < 0 \tag{41}$$

The negative relationship between growth and the real wage rate, instead, disappears. Equations (30)–(35) generate the so-called 'paradox of costs', according to which an increase in costs, in the form of a higher wage rate, implies higher profits and growth rates (see Rowthorn, 1981, p. 18 and Lavoie, 1992, p. 307). By differentiating expressions (38) and (39) with respect to w_π, one obtains

$$\frac{d\overline{r}_k}{dw_\pi} = \frac{\gamma_0\gamma_2 a_k a_l}{[(1-w_\pi a_l)(s_c-\gamma_1)-\gamma_2 a_k]^2} > 0 \tag{42}$$

$$\frac{d\overline{g}}{dw_\pi} = \frac{s_c\gamma_0\gamma_2 a_k a_l}{[(1-w_\pi a_l)(s_c-\gamma_1)-\gamma_2 a_k]^2} > 0 \tag{43}$$

The paradox of costs is caused by the fact that investment expenditures are more sensitive to changes in effective demand (reflected by the degree of capacity utilisation) induced by changes in distribution (reflected by the wage share) than to changes in costs induced by changes in the wage rate (and in the profit margin).

The analytical condition indicating when the paradox of costs occurs is given by the value of the elasticity $\xi(u, \pi) < -1$. This elasticity measures the sensitivity of effective demand to changes in distribution. From (36), in fact, the inequalities $dr_k/d\pi < 0$ and $dg/d\pi < 0$ (and, therefore, $dr_k/dw_\pi > 0$ and $dg/dw_\pi > 0$) imply $\xi(u, \pi) < -1$. For the model (30)–(35), this condition always holds since, from (37),

$$\xi(\overline{u}, \pi) = -\frac{\pi(s_c-\gamma_1)}{\pi(s_c-\gamma_1)-\gamma_2 a_k} < -1$$

Note finally that, when the wage rate exceeds the value

$$w_\pi > 1 - \frac{\gamma_2 a_k}{s_c - \gamma_1},$$

the equilibrium solution $\overline{u}$ does not satisfy the condition $u \leq 1$ and the Kaleckian analysis becomes overdetermined. When the constraint $u = 1$ is binding, firms cannot expand production to accommodate further rises in demand. The disequilibrium between demand and supply, $i > s > \overline{g}a_k/\overline{u}$, persists unless prices and profit margins rise and the wage share falls (see Rowthorn, 1981, p. 10). The neo-Keynesian adjustment mechanism is thus restored.

Moving on from the relationship between the rate of profits and the degree of capacity utilisation (36), $r_k = \pi u/a_k$, Bhaduri and Marglin (1990) amended the Kaleckian theory taking into account that investment reacts differently to similar changes in profitability. In particular, at the same rate of profit investment decisions differ when profit margins are low and capacity utilisation high and profit margins are high and capacity utilisation low. Firms may not be willing to expand further productive capacity when excess capacity is already extensive. Consequently, equation (34) has to be replaced by the following

$$\frac{i}{k} = \gamma_0 + \gamma_1\pi + \gamma_2 u \tag{44}$$

The solutions of the model (30)–(33), (35) and (44), considering that $\pi = 1 - wa_l$, are

$$\overline{u} = \frac{[\gamma_0 + \gamma_1(1-w_\pi a_l)]a_k}{s_c(1-w_\pi a_l)-\gamma_2 a_k} \tag{45}$$

$$\overline{r}_k = \frac{(1-w_\pi a_l)[\gamma_0+\gamma_1(1-w_\pi a_l)]}{s_c(1-w_\pi a_l)-\gamma_2 a_k} \tag{46}$$

$$\overline{g} = \frac{s_c(1-w_\pi a_l)[\gamma_0+\gamma_1(1-w_\pi a_l)]}{s_c(1-w_\pi a_l)-\gamma_2 a_k} \tag{47}$$

By differentiating expressions (46) and (47) with respect to w_π, one obtains

$$\frac{d\overline{r}_k}{dw_\pi} = \frac{[\gamma_2\overline{u}a_k - \gamma_1(1-w_\pi a_l)a_l]a_l}{s_c(1-w_\pi a_l)-\gamma_2 a_k} \tag{48}$$

$$\frac{d\overline{g}}{dw_\pi} = \frac{s_c[\gamma_2\overline{u}a_k - \gamma_1(1-w_\pi a_l)a_l]a_l}{s_c(1-w_\pi a_l)-\gamma_2 a_k} \tag{49}$$

The sign of the derivatives (48) and (49) depends on the parameters of the model. It follows that the model modified with the investment function (44) is able to generate two alternative growth regimes. A wage-led growth regime, characterised by $d\overline{r}_k/dw_\pi > 0$ and $d\overline{g}/dw_\pi > 0$ ($d\overline{r}_k/d\pi < 0$ and $d\overline{g}/d\pi < 0$), prevails when $\overline{u} > \gamma_1\pi a_l/\gamma_2 a_k$. The wage-led regime is characterised by great responsiveness of effective demand to changes in distribution, $\xi(\overline{u}, \pi) < -1$. The overall effect of an increase in the wage rate on growth is positive because the positive effect of demand (induced by the distribution in favour of workers) is greater than the negative effect of higher costs (generated by the increased wage rate or decreased profit margin). The paradox of costs holds. Conversely, a profit-led growth regime, characterised by $d\overline{r}_k/d\pi > 0$ and $d\overline{g}/d\pi > 0$ ($d\overline{r}_k/dw_\pi < 0$ and $d\overline{g}/dw_\pi < 0$), prevails when $\overline{u} < \gamma_1 \pi a_l/\gamma_2 a_k$. The profit-led regime is characterised by little responsiveness of effective demand to changes in distribution $\xi(\overline{u}, \pi) > -1$. Growth is enhanced by increases in the profit margin because the negative effect of changes in the wage share on demand is more than compensated by the inducement to invest caused by lower costs (lower wage rates). The negative relationship between w and r_k and g holds as in the neo-Keynesian model.

A recent attempt has been made to develop an approach (labelled neo Ricardian) to investment-led growth in line with the Classical theory of prices and distribution (see Vianello, 1985, 1989, 1996; Ciccone, 1986, 1987; Committeri, 1986, 1987; Kurz, 1986, 1992; Garegnani, 1992; Serrano, 1995; Trezzini 1995, 1998; Garegnani and Palumbo, 1998; Ciampalini and Vianello, 2000; Park, 2000; and Barbosa–Filho, 2000). In this approach the 'normal' income distribution, that is, the distribution corresponding to the degree of capacity utilisation desired by entrepreneurs (which is also labelled 'normal'),[28] is determined by conventional or institutional factors.[29] Moreover, the rate of growth of demand may affect investment decisions, as a result of firms' constant attempts to match productive capacity to expected demand. This feature is not explicitly taken into account in neo-Keynesian and Kaleckian analyses. Neo Ricardians also object that the Kaleckian approach has no adjustment mechanism between the current and normal degree of capacity utilisation.[30] However, they allow that these two magnitudes may differ for long periods of time.[31]

An attempt to clarify the neo Ricardian position is made by introducing the following equations derived from expressions (13)–(19) by assuming an endogenous degree of capacity utilisation, $u \le 1$; an exogenous income distribution, $w = w_\omega$; and disregarding the role of expected profitability in the investment function:[32]

$$1 = wa_l + r_k k \tag{50}$$

$$\frac{1}{k} = \min\left(\frac{l}{a_l}, \frac{1}{a_k}\right) \tag{51}$$

$$u = \frac{a_k}{k} \tag{52}$$

$$w = w_\omega \tag{53}$$

$$s = s_c r_k k \tag{54}$$

$$\frac{i}{k} = u - 1 + gu \tag{55}$$

$$s = i \tag{56}$$

Equation (53) assigns a conventional nature to the wage rate. Unlike the neo-Keynesian analysis, exemplified by equation (24), normal distribution, and in particular the normal rate of profits, is independent of accumulation. $r_n = 1/a_k - w_\omega(a_l/a_k)$ represents the normal rate of profits.[33] According to equation (55), investment expenditure is driven by an accelerator mechanism. The latter involves the entrepreneur's attempt to adjust productive capacity towards the planned degree (here corresponding to full capacity) and to install capacity to adjust to (expected) demand growth.

From (50)–(56), by imposing the equilibrium growth condition $u = 1$, one obtains the solutions:

$$\overline{r}_k = r_n \tag{57}$$

$$\overline{g} = s_c r_n \tag{58}$$

According to expressions (57) and (58), in equilibrium, the rate of profits coincides with its normal value and the rate of growth is governed by that level of saving, $s_c r_n$, which corresponds to normal capacity utilisation or 'capacity saving'. From this analysis it follows that, along the equilibrium path, effective demand does not affect growth.

To re-assign a role to demand the neo Ricardian literature has taken two routes. The first introduces in the equilibrium condition of the commodity market a component of demand that is independent of the level of income and its rate of change (Serrano, 1995; Park, 2000; and Barbosa–Filho, 2000).[34] The second abandons the use of equilibrium growth analysis and suggests the adoption of empirical and historical analyses, which are case-specific, in order to identify the influence of the various components of demand in different historical phases (see Garegnani, 1992; Ciampalini and Vianello, 2000; and, for an example of historical analyses, Garegnani and Palumbo, 1992).

6.5. THE INFLUENCE OF THE EXTERNAL COMPONENT OF AGGREGATE DEMAND

The analysis of the influence of the external components of demand is mainly based on the contributions of Harrod, Kaldor and Thirlwall, which point out that the rate of growth of an open economy may be constrained by its trade performance. Some insights into the role of external demand can already be found however in Keynes's writings on the British return to gold. In *The Economic Consequences of Mr. Churchill* (1925), Keynes claimed that the return to the pre-war parity would have had a negative influence on the British trade, making a sharp reduction of money wages necessary to restore the competitiveness of the national industry on overseas markets. The wage adjustment, however, would not have been painless: in the absence of a fall in the cost of living, workers' resistance to wage reductions had to be overcome 'by intensifying unemployment without limits' (Keynes, 1925, pp. 211 and 218).

At the time, the theory of international trade was dominated by 'classical' thinking, according to which the balance of payments automatically adjusts through gold flows and consequent relative price movements: countries experiencing a trade deficit would lose gold, causing an internal price deflation which would induce a rise in exports and a fall in imports such as to restore equilibrium. According to Keynes, however, gold flows may fail to restore the balance of payments equilibrium if wages and prices react slowly to changes in the quantity of money: in these cases, the 'classical' mechanism would not work, and interest rate adjustments have to come into play to ensure capital inflows sufficient to compensate for the trade deficit, discouraging capital accumulation and slackening economic activity.

In the following years, Keynes restated this view on various occasions. In the evidence addressed to the Macmillan Committee, he went so far as to advocate protectionism as a remedy against recession, a provocative suggestion in a *laissez-faire* oriented environment (Keynes, 1929, pp. 113–7). The proposal testifies to the relevance Keynes attributed to the constraint that the balance of payments can set to domestic prosperity. In his view, as long as monetary policy was sacrificed to the achievement of external equilibrium, Britain was inevitably condemned to stagnation (Keynes, 1929, pp. 56–7). To 'release' monetary policy from this task the British competitive performance in overseas markets had to be improved. This view also emerges in the *General Theory*.

> In an economy subject to money contracts and customs more or less fixed over an appreciable period of time, where the quantity of domestic circulation and the domestic rate of interest are primarily determined by the balance of payments, …, there is no orthodox means open to the authorities for countering unemployment at home except by struggling for an export surplus and an import of the monetary metal at the expense of their neighbours (Keynes, 1936, p. 348).

The idea that the trade performance of a country may affect its level of activity was restated by Harrod in his 1933 *International Economics.* Like Keynes, Harrod analyzed the case of an economy with sticky wages, where

the gold outflows caused by a trade deficit cannot affect relative prices, so that the 'classical' adjustment process does not work. In this case, the gold outflows would cause 'real' effects, and a poor trade performance may therefore become a constraint to domestic activity and employment (Harrod, 1933, pp. 118 and 125). This view is formally depicted through the so-called 'foreign trade multiplier' (pp. 119–23), that is a causal relationship going from exports to domestic output. Consider an economy with no Government sector and no saving and investment. In this case, income,Y, is spent either on home-made consumption goods, C, and imports, M:

$$Y = C + M \tag{59}$$

National income is equal to the sale of domestic goods at home, C, and exports, X,:

$$Y = C + X \tag{60}$$

If the country spends on imported commodities a stable fraction μ of its income,

$$M = \mu Y \tag{61}$$

substituting (61) in (60) and equating (59) and (60), we get:

$$Y = \frac{1}{\mu} X \tag{62}$$

The link with Keynes' insights into the influence of international trade on domestic prosperity is straightforward: when deterioration of the trade performance of a country, whether a reduction of exports or an increase in the import propensity, occurs, the commodity market equilibrium is restored through a reduction of output. Thus, the country's trade performance may constrain economic activity and employment.

Harrod's analysis of the dynamic adjustment of output following an external shock also reflects Keynes' line of reasoning: in the case of a current account disequilibrium, the gold outflows would cause pressures on interest rates, thus affecting investment in fixed and working capital and giving rise to changes in domestic output (Harrod, 1933, pp. 135–7).

Harrod noted that, under the simplified assumptions of the model, the commodity market equilibrium automatically implies $X = M$ (Harrod, 1933, p. 120). He also clarified that the relationship between foreign trade performance and domestic ouput still holds in a more general model taking into account saving and investment, even if in this case the output adjustments may no longer be sufficient to assure balanced trade.

Other contributions to the study of the role of the external component of aggregate demand in growth theories can be found in the 1960s with Kaldor's work on growth rate differentials, where this analysis was intertwined with that of cumulative causation.[35] In these works, which had a great impact on development studies and on the subsequent birth of the 'evolutionary literature',[36] Kaldor claimed that orthodox theory fails to explain the divergence in growth rates among economies, which 'are largely accounted for by differences in the rates of growth of productivity' (Kaldor, 1966, p. 104). The latter, in turn, are mainly due to the economies of scale occurring within the industrial sector, whose rate of growth shows an 'extraordinarily close correlation' (Kaldor, 1978a, p. XVIII) with the rate of growth of GDP and productivity.

In order to describe the actual performance of the economies, Kaldor (1966; 1967; 1970; 1972) used the notion of 'circular and cumulative causation', introduced by Myrdal (1957), considering the dynamics of the industrial sector as the 'engine of growth'. Following Young, Kaldor (1966 and 1967) described growth as a process generated by the interaction between demand and supply: the rate of growth is positively related to the ability of supply to accommodate variations in demand and to the reaction of demand to changes in supply. Moreover, he clarified that economies move through different stages of economic development. In an early stage, the demand for consumption goods plays the leading role in the growth process. In the later stages, the leading forces are, respectively, the export of consumption goods, the demand for capital goods, and, finally, the export of capital goods (Kaldor, 1966, pp. 112–4).

In his subsequent essays, Kaldor underlined other aspects of the growth process. In 1970 he examined how growth depends on the rate of change of exports, by applying Hicks' (1950) 'super-multiplier' to an open economy and considering exports as the leading force, and consumption and investment as induced components. The rate of growth of exports, in turn, was assumed to depend on an external cause, the world rate of growth of demand, and on a domestic cause, the rate of change of production costs. An increase in world demand raises exports and domestic production through the super-multiplier. Increasing returns in the export sector reduces costs, unless a proportional rise in wages occurs. The reduction in costs further increases exports, setting up a cumulative process, which tends to broaden the gaps with other regions.[37]

For Kaldor, therefore, the demand coming from the foreign sector plays a primary role in setting in motion the growth process, while the domestic sources of demand mainly influence the competitiveness of the economy and the intensity with which the external stimulus is transmitted to the rate of growth.

In 1975 Dixon and Thirlwall tried to embody in a formal model the view presented by Kaldor in his 1970 article. According to them, the working of the growth process in an open economy may be so depicted:

$$g = \gamma \hat{x} \tag{63}$$

$$\hat{x} = \eta_x \left(\hat{p} - \hat{p}_f - \hat{e}\right) + \varepsilon_x g_f \tag{64}$$

$$\hat{p} = \hat{w} - \hat{a}_l + \hat{\pi} \tag{65}$$

$$\hat{a}_l = a_0 + \lambda g \tag{66}$$

where g is the rate of growth of the economy, $\hat{x}$ the rate of growth of exports, $\hat{p}$, $\hat{p}_f$ and $\hat{e}$ are rates of change of domestic prices, foreign prices and exchange rates respectively, g_f is the rate of growth of world income, $\hat{w}$, $\hat{a}_l$ and $\hat{\pi}$ are rates of change of wages, labour productivity and mark-up factor respectively.

Equation (63) specifies Kaldor's idea that the rate of growth of the economy is directly related to the growth of exports.[38] Equation (64) is the dynamic formulation of a conventional multiplicative export function relating the rate of growth of exports to the rates of change of relative prices and world income, with η_x and ε_x being constant price and income elasticities. Equation (65) describes the rate of change of domestic prices as depending on changes in the unit labour costs and on changes in the mark-up factor. Finally, equation (66) describes the relation between the rate of change of productivity and the rate of growth of output known in the literature as the *Verdoorn's Law*.[39]

The equilibrium solution of equations (63)–(66) is

$$g = \frac{\gamma[\eta_x(\hat{w} - a_0 + \hat{\pi} - \hat{p}_f - \hat{e}) + \varepsilon_x g_f]}{1 + \gamma\eta_x\lambda} \tag{67}$$

Dixon and Thirlwall (1975) also presented the model in terms of finite difference equations, deriving equation (67) as the steady growth solution.[40] This equation can be used to describe the evolution of the rates of growth of different countries or of different regions within the same country. If one assumes a given mark-up in each region and given and equal values of $\hat{p}_f$, g_f, and $\hat{w}$ in all regions,[41] the differences in the rates of growth depend on the regional values of λ, γ, η_x, ε_x, and a_0.

Owing to its 'aggregate' structure, the model (63)–(66) neglects the role of the sectoral composition of the economy and, therefore, it does not adequately depict the richness of Kaldor's views on growth, based on the idea that the productive structure affects the overall rate of growth of productivity. Yet, the relevance of these 'composition effects' may be easily taken into account by analysing how the sectoral composition of the economy affects the parameters of the model.

As to λ, Kaldor (1971) argued that it mainly depends on the composition of demand and on the weight of the capital goods sector in the productive structure. High investments and a large capital goods sector enhance productivity and the competitive performance of the economy in the world markets.[42] According to Kaldor (1966; 1967; 1971), the influence of the composition of demand on productivity is due to the presence of variable returns in the different sectors of the economy. The intensity of the effect on productivity thus crucially depends on the sectors towards which the demand for consumption and investment is directed, since increasing returns mainly occur in the capital goods sector. Moreover, the extent to which this sector is able to accommodate demand is also important. High quotas of investment to output and of the capital goods sector in the productive structure enhance productivity changes, which, in turn, improve the international performance of the economy setting up and intensifying cumulative processes.

Kaldor (1971) referred to the role of composition of demand on long-term growth in his policy analyses too. He distinguished between the concepts of 'consumption-led' and 'export-led' growth, arguing that the latter is more desirable than the former: consumption-led growth tends to have negative long-run effects on productivity, since it tends to raise the weight of non-increasing return sectors in the productive structure of the economy. This tends to worsen the international performance of the economy. Hence, as stated in section 6.3 above, Kaldor claimed that Government intervention should avoid the use of fiscal policy to increase the rate of growth and reduce unemployment. By making growth more dependent on the demand for consumption, this policy generates the undesired consequences previously recalled. In this case, he said, the authorities should intervene on the exchange rate, rather than through fiscal measures.[43]

Kaldor's writings also hint at the factors affecting γ, which depends on the quotas and elasticities of the various components of domestic demand to the net output of the economy.[44] The elasticity of the demand for consumption is influenced by productivity growth through the introduction of new products of large consumption (Kaldor, 1966, p. 113; 1981, p. 603; and Rowthorn, 1975, p. 899). When this occurs a higher value of γ and a more intense effect of a given rate of growth of exports come about. For Kaldor (1971) tax reduction too has a positive influence on γ, through its effect on consumption.[45] Yet, any stimulus to the latter variable has long-run negative consequences, as stated above, since it makes the growth process consumption-led. Finally, the elasticity of imports depends on the degree of

coincidence between the composition of demand and the productive structure of the economy. In 1966 Kaldor related the degree of coincidence of the productive structure to demand to the stage of development reached by a country. The more a country can rely on a large capital goods sector, the lower will be the elasticity of imports, the higher the value of γ and the more stimulating the effect of a given rate of change of exports. A country that has reached a stage of development which allows it to be a net exporter of capital good can enjoy 'explosive growth', since 'a fast rate of growth of external demand for the products of the 'heavy industries' is combined with the self–generated growth of demand caused by their own expansion' (Kaldor, 1966, p. 114).

An important and controversial issue concerns the factors affecting η_x and ε_x. Kaldor (1971) considered price competitiveness the most important factor at work. In Kaldor (1978c) this position was abandoned, on account of the fact that the worst performing countries in terms of relative prices after the 2nd World War proved to be the best performing in terms of exports (McCombie and Thirlwall, 1994, pp. 262–300). Kaldor (1981) then concluded that the rate of growth of exports mainly depends on income elasticity, which in turn depends on the innovative capacity of a country, that is, the capacity of a country to differentiate its products. This innovative capacity gives the economy a privileged position in foreign markets.

In their 1975 paper, Dixon and Thirlwall also tested their model on United Kingdom data, but the model gave rise to unsatisfactory approximation between fitted and actual values over the period 1951–66, since higher than actual growth rates were systematically predicted. According to Thirlwall (1998, p. 194) this discrepancy could be explained by the neglect of the balance-of-payments constraint, in that period a severe hurdle to Britain's growth performance. To make up for this failure, in 1979 Thirlwall worked out an analytical model incorporating the external equilibrium condition, described by the following equation:

$$p\,X + F = p_f\,M\,e \tag{68}$$

where p is the export price index, p_f the import price index, e the exchange rate and F the value of net capital flows measured in domestic currency. Expressing (68) in terms of rates of change, we get:

$$\theta(\hat{p}+\hat{x})+(1-\theta)\hat{f} = \hat{p}_f + \hat{m} + \hat{e} \tag{69}$$

where $\hat{m}$ and $\hat{f}$ denote respectively the rate of growth of imports and the rate of change of net capital flows, while θ and $(1 - \theta)$ are respectively the value of exports and capital inflows as a percentage of imports. If we specify the demand for imports and exports through the conventional multiplicative functions with constant elasticities, we may express the rate of change of exports through equation (64) and the rate of change of imports by:

$$\hat{m} = \eta_m(\hat{p} - \hat{p}_f - \hat{e}) + \varepsilon_m g \tag{70}$$

where η_m and η_m are price and income elasticities respectively. Substituting (64) and (70) in (69) and rearranging, we get:

$$g_B = \frac{\theta\varepsilon_x g_f + (1-\theta)(\hat{f} - \hat{p}) + (1+\theta\eta_x + \eta_m)(\hat{p} - \hat{e} - \hat{p}_f)}{\varepsilon_m} \tag{71}$$

where g_B is the rate of growth consistent with equilibrium in the balance of payments. Basing his work on the extensive empirical evidence showing long-run stability in the terms of trade,[46] Thirlwall assumed that the contribution to growth of the price term in (71) is likely to be small. If for simplicity's sake it is assumed to be zero, equation (71) reduces to:

$$g_B = \frac{\theta\varepsilon_x g_f + (1-\theta)(\hat{f} - \hat{p})}{\varepsilon_m} \tag{72}$$

If we also assume that a country cannot finance its trade deficit through capital inflows for a considerable length of time, the long-run equilibrium requires that $\theta = 1$ (McCombie, 1998, pp. 229–32). Equation (72) changes into

$$g_B = \frac{\varepsilon_x}{\varepsilon_m} g_f \tag{73}$$

which represents the dynamic version of Harrod's foreign trade multiplier. The economic meaning of equation (73) is that a poor trade performance constrains a country to grow at a slower pace than that allowed by the growth of internal demand and by resource availability. If $g > g_f$, imports would grow quicker than exports, worsening the country's trade account and forcing policy-makers to intervene. When for various reasons (*real wage-resistance* and subsequent transmission of exchange rate variations on domestic prices, product differentiation leading to small price elasticity of demand for tradable goods, etc.) exchange rate devaluations prove ineffective, the balance of payments adjustment takes place through internal demand deflation, which slackens the pace of growth (Thirlwall, 1979, pp. 279–80). Analogously, if $g < g_f$ and the country is able to expand internal demand, the pressure of demand upon productive capacity may raise the

capacity growth rate up to the ceiling represented by equation (73) According to this approach, capital and labour availability does not constrain growth, being to a large extent 'endogenous' to the economic system.[47]

The relevance of equation (73) lies in the fact that it supplies a simple and attractive explanation of why growth rates differ among countries. An increase in world income generates a rate of growth that depends on the value of each country's $\varepsilon_x/\varepsilon_m$ ratio. Since there are significant international differences in this ratio (Houthakker and Magee, 1969), the same increase in the world income gives rise to different growth rates among countries.

A relevant question, to which this strand of literature has not yet given a conclusive answer, is what determines the $\varepsilon_x/\varepsilon_m$ ratio. In some contributions, Thirlwall (1979, p. 286 and 1991, p. 26) claims that the differences in this ratio mainly reflect those in the patterns of productive specialization. This way of interpreting the dynamic foreign trade multiplier has striking implications for the theory of uneven development. For example, assume a simplified world where some countries only produce manufactured goods and others only produce primary goods. As the income elasticity of the demand for manufactured goods, due to Engels' Law, is higher than income elasticity of the demand for primary goods, it would be $\varepsilon_x/\varepsilon_m > 1$ for countries producing manufactured goods and $\varepsilon_x/\varepsilon_m < 1$ for those producing primary goods. According to this view, therefore, the pattern of specialisation is the source of a process of cumulative divergence in GDP levels: countries producing primary goods would be unable to grow at the same rate as those producing manufactured goods, owing to their tighter balance-of-payments constraint.

Although attractive, this way of interpreting the foreign trade multipliers has been poorly supported on empirical grounds,[48] inducing Thirlwall to return to the topic and clarify that, for industrial countries, income elasticities must also be made to depend on the supply characteristics of the goods produced, such as their technical sophistication and quality (see Thirlwall, 1991, p. 28 and 1998, p. 187). With this revision, the 'cumulative divergence' view rooted in the post-Keynesian tradition may be extended even to growth differentials among industrial countries: in Thirlwall's view, indeed, an initial discrepancy in growth rates sets in motion the negative feedback mechanisms associated with Verdoorn's Law, which 'will tend to perpetuate initial differences in income elasticities associated with "inferior" productive structures on the one hand and "superior" industrial structures on the other' (Thirlwall, 1991, p. 27).[49]

Thirlwall's 1979 analysis has been subsequently extended to take into account the role of international capital flows. Thirlwall and Hussain (1982) used equation (72), instead of (73), to capture the experience of some developing countries running persistent current account deficits, financed by foreign investment. In some more recent contributions (Moreno Brid, 1998–99, McCombie and Thirlwall, 1999), however, the use of equation (72) has been considered inappropriate for a steady-state analysis without imposing any restriction on the evolution path of foreign capital inflows, as the lack of this restriction may generate a path of foreign debt unsustainable in the long run. According to Moreno Brid (1998–99), international credit institutions impose on developing countries borrowing restrictions based on some index of their expected ability to repay the foreign loans. He therefore proposes a different specification for the balance-of-payments constraint based on the requirement of a constant ratio between the current account deficit and the GDP, interpreted as a measure of a country's creditworthiness. When this restriction is added to the model, the dynamic foreign trade multiplier may assume a value higher or lower than the standard one, depending on the initial current account position of the country concerned. This revision has considerable implications for empirical analysis, clarifying that estimates of the $\varepsilon_x/\varepsilon_m$ ratio may be significantly biased if they do not take into account the countries' initial export/import ratio.

To sum up, the balance-of-payments constraint approach provides some important insights into the analysis of the relationship between external demand and growth. While on theoretical grounds the relevance of the cumulative causation mechanism embodied in the model (63)–(66) cannot be denied, the empirical evidence seems to show that the simpler formula described by equation (73) suffices to capture the main 'stylised facts' relating to growth.[50] As the analysis of the factors affecting the $\varepsilon_x/\varepsilon_m$ ratio seems to suggest, however, the balance-of-payments constraint approach does not obscure the peculiar role played by the interaction between 'external' and 'internal' factors underlined by Kaldor in his writings.

6.6. CONCLUSIONS

Harrod's seminal work on growth theory was conceived as an attempt to extend Keynes's analysis. It moved from the Keynesian ideas that the economic system does not tend necessarily to full employment and that aggregate demand may affect the rate of growth of the economy. In subsequent years, Keynesian economists developed this approach along several lines, focussing on the different components of aggregate demand and on their role in the growth process, by using several descriptive and analytical methods. As stated above, this multiplicity of ideas and analyses shows, according to some authors, the fertility of this line of thought. Conversely, an external observer may judge the lack of a unified framework a weakness, considering the Keynesian literature a disorderly set. By

reconstructing the content of a Keynesian approach to growth and describing the lines of development that have historically emerged, this paper has tried to underline the wealth of this tradition. At the same time, it has sought to outline the existence of some unifying elements which, while preserving the diversity of ideas and analyses, reduces the risk of interpreting the Keynesian literature as a disorganised set.

NOTES

1. The model proposed by Solow (1956) describes the neoclassical theory of growth. For the classical tradition one can refer to the analyses proposed by Pasinetti (1960) and by Samuelson (1978). The analyses presented by Barro and Sala-i-Martin (1995) give the main elements of the New Growth Theories.
2. This multiplicity of ideas and analyses is, according to some authors (e.g. Dow, 1985; Hamouda and Harcourt, 1989; and Chick, 1995), a great merit of the Keynesian literature, since it adds to the richness of this line of thought.
3. See Rochon (1999, pp. 64–9) for a collection of these criticisms against Keynesian economics, raised by authors like Solow, Backhouse, Dornbusch, Fisher, Felderer and Homburg.
4. 'During the twenties many of us were deeply interested in Keynes's advocacy of measures to promote fuller employment' (Harrod, 1967, p. 316).
5. Harrod (1951, ch. IX, par. 3) recalls however Keynes's article in the *Nation*, May 24, where the Cambridge economist presented for the first time his proposals for public works.
6. Phelps Brown (1980, p. 19) points out that since 1932, Harrod wrote several letters to *The Times*, in favour of Keynes's proposals.
7. See Harrod (1948, p. 40; 1964, pp. 903 and 905–6). The similarity between Harrod's and Ramsey's analysis of saving is underlined by Asimakopulos and Weldon (1965, pp. 66). Harrod (1973, p. 20) also clarifies that 'what each person chooses in regard to saving is governed by various institutional arrangements, which differ from country to country and from time to time. There is the question of what the State will provide for future contingencies – old age, ill health, unemployment, etc. – by current transfer payments as and when they arise. The more ground that the State covers, the less will the individual feel it incumbent to provide for himself by saving. Personal saving will also be affected by the degree the education of one's children is subvented by the public authorities'.
8. According to Harrod (1939, p. 264), the warranted rate is the rate that, if it occurs, leaves producers satisfied, in the sense that for them 'stock in hand and equipment available will be exactly at the level they would wish to have them'.
9. Harrod (1948, p. 83) points out that his analysis of the warranted rate assumes the rate of interest constant. He referred to the realism of Keynes's view on the behaviour of the interest rate (pp. 64–5), agreeing that this rate may be rigid (pp. 56–7) and unable to decrease in such a way as to lead to full employment (pp. 70–1; 83–4; 97; 99).
10. See Harrod (1939, pp. 258, 259 and 276). On page 276, in particular, Harrod explicitly referred to an inverse relationship between *k* and *r*. In the Thirties the neoclassical assumption of decreasing marginal returns was generally accepted. Sraffa's critique of the neoclassical theory of capital had not yet been elaborated. (See Panico, 1998, p. 177, fn. 55; and Panico, 2001, pp. 300 and 308–9 fn. 59, 60 and 61). As is well known, it was published in 1960 and discussed at length in the following decade.

11. Dealing with his analysis of the equilibrium warranted path, Harrod claimed: 'I know of no alternative formulation, in the world of modern economic theory, of any dynamic principle of comparable generality. We must start with some generality however imperfect. We shall never go ahead if we remain in a world of trivialities or fine points. It is useless to refine and refine when there are no basic ideas present at all' (Harrod, 1948, pp. 80–1).
12. As to the 'knife-edge problem' Harrod stated: 'Nothing that I have ever written (or said) justifies this description of my view' (Harrod, 1973, p. 31; but see also pp. 31–45).
13. See Harrod (1948, pp. 132–3, 137–8 and 144; 1960, pp. 278–9, 283 and 285; 1964, pp. 910–13; 1973, pp. 68, 78, 80, 102). It should be noted too that, after 1960, Harrod thought that the major influence of the interest rate on investment is through the availability of finance, owing to the fact that the credit markets are imperfect (information are asymmetrically distributed) and tend to react to the shortage or availability of credit (see Harrod, 1960, pp. 278–9 and 292; 1964, pp. 912–13; 1973, pp. 44, 61, 179).
14. 'Sustained low interest will presumably in the long run reduce the normal profit rate' (Harrod, 1973, p. 111). And again: 'If the market rate of interest rises considerably and stays up for a substantial period, ... that may cause firms to increase the mark-up' (p. 44; see also, p. 78).
15. Harrod (1964, p. 908) gave a somewhat different account of this point: 'In the concluding pages of my first "Essay" I did recognise that there were two distinct problems of policy, namely: (*i*) the short-term one of preventing deviations from a steady growth rate, and (*ii*) the long-term one of bringing the warranted rate into line with the natural growth rate. I recognised that, if the warranted rate was not equal to the natural rate – and there is no reason why it should be – difficulties would inevitably arise. Thus, policy was required to bring them together. My remarks on this subject were admittedly very sketchy. I suggested that the long-term interest rate might be used to make the warranted rate adhere more closely to the natural rate, while "public works" (nowadays "fiscal policy") and the short-term rate of interest should be used to deal with short-term deviations. All this was very loose. The existence of the double problem was, however, recognised'.
16. Some recent contributions to the New Growth Theories consider, instead, the influence of Government intervention on growth, be it a change in taxation or in expenditure, through its effect on the propensity to save and on the capital–output ratio (see Barro, 1990).
17. In his seminal contribution Kaldor (1955–56, p. 98) explicitly recognised the need to deal with the State in the analysis of steady growth conditions. Yet, like other authors, he failed to do so in most of his later work.
18. According to Kaldor (1958, pp. 136–7), the drawback of this solution is that in times of inadequate demand the Government gradually transforms the economy into one of high consumption and low investment, with the undesirable consequences on long-run growth, which will be described in Section 6 below.
19. In this debate, Pasinetti (1989a; 1989b) and Dalziel (1989; 1991a,b; 1991–92) examine the validity of the Cambridge equation by introducing into the analysis the Ricardian debt/taxation equivalence. Denicolò and Matteuzzi (1990) and Panico (1993, 1997, 1999) consider the same topic by introducing into the analysis the existence of financial assets issued by the Government. Commendatore (1994, 1999a), instead, compares the limits of validity of the dual and the Pasinetti theorem.
20. Denicolò and Matteuzzi (1990) deal with the so-called 'personal' version of the post Keynesian theory of growth and distribution. It may be noted, however, that the debate has considered different versions of the post Keynesian theory of growth and distribution: the personal version, in terms of classes, the functional version, in terms of income groups, and the institutional version, in terms of sectors of the economy (see Panico, 1997 and Commendatore, 1999a, 1999b).

21. Kaldor (1955–56) and Pasinetti (1962), instead, assume that investment is exogenous. Their models are characterised by full employment. According to some authors this assumption cannot be considered Keynesian (see Marglin, 1984a, p. 533–4 and Kurz, 1991, p. 422). In section 3 above, however, we have pointed out that for Kaldor, full employment growth can be achieved through suitable policy interventions. In the absence of government interventions, the economy does not necessarily grow at the full employment rate. Pasinetti, on the other hand, explicitly investigates the conditions of steady growth at full employment. For a survey of the subsequent developments of the neo-Keynesian theory, see Baranzini (1991) and Panico and Salvadori (1993).
22. The introduction of a non-linear form for expression (22) could generate multiple solutions, some of them unstable. This is the case of Joan Robinson's (1962) well-known 'banana diagram' which gives rise to two equilibria, one stable and one unstable.
23. Marglin (1984a, 1984b) solved this type of overdetermination by introducing in the analysis a new variable, the rate of inflation, depending on the discrepancy between s and i. According to this author, 'equilibrium can be characterised in terms of investment, saving, and conventional wages, but to do so we must abandon the static characterisation of equilibrium in favour of a dynamic one. Using the disequilibrium dynamics of the two systems, we can synthesise Marxian and Keynesian insights into a just-determined model in which investment, saving, and the conventional wage jointly determine equilibrium' (Marglin, 1984b, pp. 129–30).
24. Dutt (1987; 1990) presented a more refined resolution mechanism of conflicting claims between firms and workers which could generate a value of the wage rate between w_π and w_ω.
25. In the Kaleckian literature these coefficients are not univocally interpreted. According to Dutt (1984, p. 28), γ_0 and γ_1 accounts for the (constant) entrepreneurs' desired degree of capacity utilisation. Lavoie (1992, 1995), instead, interpreted γ_0 as firms' expected rate of growth of sales, which is not necessarily constant.
26. These are the same reasons invoked by Joan Robinson (1962). See above.
27. According to Steindl (1952), firms plan a reserve of excess capacity facing uncertainty. This is to avoid the permanent loss of market share owing to the temporary inability to fulfil unexpected demand. Other reasons, invoked by the literature to justify firms' planned excess capacity, are: (*i*) seasonal fluctuations of demand; (*ii*) expected growth in demand; (*iii*) costly use of overtime work and night shifts or shifts involving unordinary hours or days; (*vi*) indivisibility of plants and equipment. For a short review on this argument, see Lavoie (1992, pp. 124–6).
28. The normal degree of capacity utilisation, u_n, is 'the degree of utilisation of capacity desired by entrepreneurs, and on which, therefore, they base their investment decisions about the size of a new plant relative to the output they expect to produce' (Garegnani, 1992, p. 55).
29. In particular, income distribution can be determined either by referring to some 'conventional standard of life', which affects the wage rate, or, alternatively, by the level of the money interest rates, which affects the rate of profits, as suggested by Sraffa (1960, p. 33) and envisaged by Vianello (1996).
30. On the absence of an adjusting mechanism between u and u_n, Committeri warned that if 'the "equilibrium" utilisation degree does not coincide with its normal level, and hence producers' expectations are not being confirmed by experience … as the economy moves away from the steady path, the model has nothing to say about the long-run tendencies of capital accumulation' (Committeri, 1986, p. 175). See also Ciampalini and Vianello (2000).
31. According to Garegnani (1992, p. 59), 'the entrepreneurs will certainly attempt to bring about, through investment, a capacity which can be used at the desired level. And the degree of their success will depend on how well they will be able to forecast the outputs which it

will be convenient for them to produce. But given the initial arbitrary level of capacity that success will show only in shifting, so to speak, backward in time the deviation of the utilization of capacity from the desired level. Even correct foresight of future output will not eliminate average utilization of capacity at levels other than the desired one'.

32. Neo-Ricardians consider the normal rate of profits, r_n, a more suitable variable than the current rate of profits, r, to capture the role of expected profitability in investment decisions. See on this point Vianello (1996, p. 114).
33. We assume, for simplicity, that normal and full capacity utilisation coincide, $u_n = 1$.
34. The independent component of aggregate demand can come from any sector of the economy. Notice that this analysis only shows that effective demand can affect the adjustment path towards equilibrium even if along this path $u = 1$ (see Park, 2000, pp. 11–16 and Barbosa-Filho, 2000, p. 31). As to the conclusion that equilibrium growth is governed by capacity saving, Park (2000, p. 8) and Barbosa-Filho (2000, p. 31) showed the existence of two solutions of this analysis. The first, which is locally stable, confirms that growth is governed by capacity saving. The second, which is unstable, implies that income grows at the same rate as the independent component of demand, if the latter has certain properties.
35. For an analysis of Kaldor's views on growth and cumulative causation, see Thirlwall (1987) and Ricoy (1987; 1998). They describe several aspects of Kaldor's position, including the role of technical progress and structural change, and his idea of growth as a path-dependent process. In what follows, we mainly focus on the role of demand in the growth process, paying less attention to other equally relevant aspects of his vision of the topic.
36. This is the literature that moves from the contributions of Nelson and Winter (1974, 1977, 1982), examined by Santangelo's essay in this volume.
37. In 1972 Kaldor further integrated Young's analysis with the Keynesian principle of effective demand, examining the role played by the demand for investment and focusing on the conditions allowing self-sustained growth. In this contribution, he argued that growth is a fragile process. In order to work it requires that several things simultaneously occur: investors must have confidence in the expansion of the markets; the credit and financial sectors have to accommodate the needs of trade; the distributive sector has to bring about price stability. According to Kaldor, after the 1930s, Government intervention secured the smooth working of the process by demand-management policies (Kaldor, 1972, p. 1252).
38. As stated above, Kaldor borrowed this relationship from Hicks' super-multiplier. Following standard notation, $I + X = S + M$ is the commodity market equilibrium condition for an open economy without public sector. If we assume that $S = sY$, $I = \kappa Y$ and $M = \mu Y$, the equilibrium level of income is given by $Y = \alpha X$, where $\alpha = 1/(s - \kappa + \mu)$ is Hicks' super-multiplier. In terms of rates of change, we get $g = \alpha(X/Y)\hat{x}$. Since $\alpha = dY/dX$ and, by definition, $\gamma = (dY/dX)(X/Y)$, the rate of change of income simply reduces to (63).
39. Dixon and Thirlwall (1975, pp. 208–10) point out that a_0 is determined by the autonomous rate of disembodied technical progress, by the autonomous rate of capital accumulation per worker and the extent to which technical progress is embodied in capital accumulation. λ is instead determined by the induced rate of disembodied technical progress, by the degree to which capital accumulation is induced by growth and the extent to which technical progress is embodied in capital accumulation.
40. The stability condition of the model is $|\gamma\eta_x\lambda| < 1$, which, in their opinion (1975, p. 208), may be plausibly assumed to hold. As a consequence, since $\eta_x < 0$, in equation (5.9) g is related positively to γ, a_0, $\hat{p}_f$, $\hat{e}$, ε_x, g_f and λ, and negatively to $\hat{w}$ and $\hat{\pi}$. The effects of variations of η_x are not determined. Notice too that recently Setterfield (1997) has presented an analysis, similar to that of Dixon and Thirlwall (1975), in order to study the movements of the economy out of equilibrium.

41. See Dixon and Thirlwall (1975, p. 209). Notice that, on the contrary, Kaldor (1966, p. 147) assumes that the differences in the rate of change of money wages of different regions do not counter-balance the reduction in costs due to the different rate of change of productivity.
42. To empirically estimate the influence of the composition of demand on productivity, Kaldor (1966) also used an expression, which differs from our equation (66) only in introducing, as an additional variable, the ratio of investment to output. His analysis showed that this variable explained the divergence of the rate of change of productivity from the trend determined by the original equation (66). It explains the residual change in productivity, not explained by increasing returns.
43. In the subsequent years, Kaldor changed this position too: 'In this respect I now feel I was mistaken. Events since 1971 have shown that the exchange rate is neither as easy to manipulate nor as rewarding in its effect on the rate of growth of net exports as I have thought' (Kaldor, 1978a, p. XXV).
44. Let $Y = D + X$, where Y is income, D is the demand for domestic products and X is exports. By definition $\gamma = \omega_x(dD + dX)/dX$, where ω_x is the ratio of exports to income. Since $(dD/dX)\omega_x = (dD/dY)(dY/dX)\omega_x = (dD/dY)\gamma$ and $\omega_x = 1 - \omega_D$, we can write $\gamma = (1 - \omega_D)/(1 - dD/dY)$. Finally, from the definition of the income elasticity of demand for domestic products ε_D, we get $\gamma = (1 - \omega_D)/(1 - \omega_D\varepsilon_D)$.
45. This view was already presented in Kaldor (1958), as stated in Section 3 above.
46. See Wilson (1976), Ball, Burns and Laury (1977). Long-run stability in the terms of trade may alternatively rely either on arbitrage or on wage-resistance forcing domestic prices to move equiproportionately to exchange rate depreciations so that $\hat{p} - \hat{p}_f - \hat{e} = 0$ (Thirlwall, 1979, p. 283).
47. According to McCombie and Thirlwall (1994, 233), there are a number of possible mechanisms through which capacity growth may adjust to demand growth: 'the encouragement to invest which would augment the capital stock and bring with it technological progress; the supply of labour may increase by the entry of the workforce of people previously outside or from abroad; the movement of factors of production from low productivity to high productivity sectors, and the ability to import more may increase capacity by making domestic resources more productive'. On this point, see also Thirlwall (1986, pp. 48–9) and McCombie (1998, pp. 238–9).
48. See McCombie (1993, p. 481), who quotes extensive empirical evidence showing that income elasticities are not related to the differing product mixes of the exports of the various countries.
49. It is worth noting that alternative ways of interpreting the foreign trade multipliers may lead to less pessimistic conclusions. Bairam (1993), for example, shows the existence of a statistically significant inverse relationship between the $\varepsilon_x/\varepsilon_m$ ratio and the stage of economic development of the country, proxied by per-capita output. Such a relationship implies that developing countries are less balance-of-payments constrained than developed countries, and therefore provides some support for the 'catching-up' hypothesis: if developing countries are able to grow quicker than developed ones, GDP levels will inevitably converge in the long-run.
50. See McCombie and Thirlwall (1994, 434). Kaldor himself (1981, p. 602) admitted the utility of the simplified model. In the same essay, Kaldor assumed that the sum of the marginal propensities to consume and invest is equal to unity. This assumption transforms Hicks' supermultiplier into Harrod's multiplier. If we also assume $\eta_x = 0$, equation (67) collapses to the dynamic foreign trade multiplier. Note that the assumption $c + \kappa = 1$ has also been used in the Cambridge Economic Policy Group model. On this point see also Targetti (1991).

7. Should the theory of endogenous growth be based on Say's law and the full employment of resources?

Fabio Petri

7.1. INTRODUCTION*

In the writings of the modern Endogenous (or New) Growth theory the term 'growth' is used to mean growth of *per capita* output, with a striking definitional change relative to a couple of decades ago. This change reflects the belief that the theory of long-run *aggregate* output growth is a solved problem, there being no doubt that long-run output growth is, with sufficient approximation, determined by the reinvestment of full-employment savings (or of the savings associated with a natural rate of unemployment or NAIRU).[1]

I will argue that, on the contrary, the problem of the determinants of the growth of aggregate output is far from solved, because there are good reasons to assign aggregate demand an extremely important autonomous role; and that Endogenous Growth theorists ought seriously to reconsider this question, because it may have relevant consequences also for the growth of per capita output, or more precisely, of labour productivity.

The structure of the paper is as follows. Section 7.2 delineates the premises for an alternative theory which makes output growth depend on the evolution of aggregate demand. This is done in the form of a criticism of an opinion to the contrary recently expressed by Robert Solow. A numerical example is presented to make the point that the flexibility of capacity utilization and the adaptability of labour supply make the growth both of output and of productive capacity depend on the growth of [the autonomous components of] aggregate demand. Section 7.3 then argues that the faster is the growth of output, the greater is the effect of all the forces which influence the growth of labour productivity in Endogenous Growth models; therefore the moment the full-employment, normal-capacity-utilization assumption is dropped and a possible autonomy of the evolution of aggregate demand is admitted, the growth rate of aggregate demand must be admitted to be an

important influence on the growth rate of per capita output. Then another possible cause appears of differences among nations in the growth rate of per capita output: the different growth rate of aggregate demand. Sections 7.4 and 7.5 survey the main reasons to question the belief that the full employment of resources, or the theory that the economy gravitates around a NAIRU, are good starting points for the theory of growth. Section 7.4 discusses the relevance of the speed with which the economy converges, if at all, toward the full employment of resources, and surveys some criticisms of the supposed tendency of the economy toward a NAIRU. Section 7.5 questions the ability of the rate of interest to bring investment into line with savings.[2]

7.2. THE POSSIBLE ROLE OF DEMAND IN OUTPUT GROWTH

7.2.1. Textbooks on growth theory such as those by Barro and Sala-i-Martin (1995) or Charles Jones (1998) present only one theory of long-run output growth: investment is determined by the savings corresponding to the full employment of resources. This is also the growth theory in many intermediate macroeconomics textbooks. Probably this is the sole view on output growth to which many of to-day's young economists have ever been introduced.

I therefore start by indicating why and how a different approach, recognizing an independent role of aggregate demand, is possible. A good way to do this is by commenting upon an interesting passage in a recent paper by Robert Solow (1997) on the state of macroeconomics. Solow, after reaffirming his faith in full-employment models (e.g. his own model) for the study of long-run growth, admits that in the short run there may be deviations from full employment, and then continues:

> One major weakness in the core of macroeconomics as I have represented it is the lack of real coupling between the short-run picture and the long-run picture. ... A more interesting question is whether a major episode in the growth of potential output can be driven from the demand side. Can demand create its own supply? The magnitudes suggest that it would be awfully difficult for a surge of aggregate demand to generate enough investment to provide the capacity necessary to accommodate it. In special circumstances it might be done, say, in an economy that has a pool of labour (rural, foreign) that it can mobilize. It might also work if strong aggregate demand can induce a rise in total factor productivity... The demand-driven growth story sounds quite implausible to me under current conditions: but it is an example of the kind of questions that needs to be asked (Solow, 1997, pp. 231–2).

Note that the *possibility* of an autonomous influence of aggregate demand on long-run growth is admitted; the scepticism as to its plausibility appears empirically motivated: but the reasons remain vague; nor, I will argue, can persuasive reasons be found.

First of all, Solow's restriction of the problem only to 'a *surge* of aggregate demand' is unwarranted. Growth theory must also explain episodes of lasting *slowdowns* of aggregate demand, and for these it seems clear that there is no impediment to a lengthy economic crisis discouraging investment and thus decreasing capacity to the lower level appropriate to the lower level of demand (and this by itself suggests a potentially very important role of aggregate demand). But Solow's thesis that 'it would be awfully difficult for a surge of aggregate demand to generate enough investment to provide the capacity necessary to accommodate it' appears to be very often wrong also for *accelerations* of aggregate demand. I show this with a numerical example, crude but I think sufficient to make the point.

7.2.2. As a preliminary to the example, let us note that an aspect of the functioning of industrial economies, which appears to be forgotten or underplayed in much current growth theory, is the *adaptability of productive capacity to demand*, due to the variability of the degree of utilization of productive capacity (Garegnani and Palumbo, 1998; Trezzini, 1995).

Firms want to maintain spare capacity because that allows them to meet expected, or possible, demand fluctuations, or expected or possible growths of demand; and because very high rates of utilization cause higher average costs since overtime-labour and night-labour are more expensive.[3] The first group of reasons for spare capacity means that firms will produce more at no extra average cost if only demand is, on average, higher than expected. Even the last reason is usually not sufficient to prevent an increase in production if the demand for the product increases, even if the price of the product remains constant, because in imperfectly competitive markets (the most frequent market form) the fear of losing market shares will make a higher degree of capacity utilization than the long-period optimal one become convenient in the short run – and this without any redistribution of income away from labour: the opposite is on the contrary the more likely case because of the increased share of overtime wages.

Such variability in capacity utilization is of course what makes the production of goods capable of that rapid adaptation to changes in the level or composition of demand, which is usually observed in reality. This confirms that the variability of capacity utilization is considerable not only downwards – what nobody would deny – but also, up to limits rarely reached, upwards. The implication of this variability is that the level of production is quite variable in response to variations in demand, not only for single industries, but also for entire sectors, in particular, for the capital-goods-producing sector, and – the moment one admits (see below) that the supply of labour is not usually fully utilised and can usually be increased (in the short period perhaps by overtime work) – also for the entire economy. (Obviously the possibility of accelerating the rate of production is due to the existence of inventories of intermediate goods, inventories which will be initially run down but will be then rapidly reconstituted by the increase itself in production.)

But if the level of production is so variable, then – concentrating now on the capital-goods sector, i.e. the sector whose production creates productive capacity – the production of productive capacity, and therefore the rate of growth of productive capacity, must be considered determined by demand, so the evolution over time of the overall productive capacity of an economy must be considered, in an analysis of growth, the result of demand, rather than a determinant of production.

7.2.3. To make such statements more concrete and also stress the cumulative effects of a higher or lower aggregate demand, let us use a numerical example. This is based on a multiplier–accelerator model of a closed economy with public expenditure and a balanced budget ($G = T$), and with depreciation and a distinction between gross and net investment.[4] The actual capital stock is K, the desired capital stock is K^*. Net investment I_N in each period t is performed to bring the capital stock at the beginning of the period, K_t, to its desired level by the end of the period, i.e. at the beginning of the next period, K^*_{t+1}. The desired capital–output ratio K^*_t/Y_t is 1, and firms are assumed to be myopic or very prudent and, when deciding investment for period t, to expect for period $t+1$ the level of gross output just observed, i.e. Y_{t-1}. Thus $K^*_{t+1} = Y_{t-1}$ and *net* investment is governed by $I_{Nt} = Y_{t-1} - K_t$, while *gross* investment I_t is equal to net investment plus depreciation: the latter is assumed to be 10% of K_t. Investment plans are realized and therefore $K_t = Y_{t-2}$. Consumption is equal, with a one-period lag, to 8/9 of after-tax income, and the state is assumed to be able every period to tax for an amount exactly equal to public expenditure, so $C_t = (8/9)(Y_{t-1} - G_{t-1})$. Output is assumed to adjust very rapidly to demand so that in every period $Y = C + I + G$. The autonomous role in aggregate demand is taken by public expenditure G. The economy is assumed to be initially stationary, with $Y = 1000$, $K = K^* = 1000$, $I = 100$, $I_N = 0$, $G = 100$, $C = 800$. Initially a zero growth rate is Harrod's warranted growth rate. Then from period 0 onwards public expenditure starts increasing at a rate of 2% per period. The following table describes the initial evolution of the economy and the 15th and 16th periods:

t	G_t	$K^*_{t+1} = Y_{t-1}$	$K_t = K^*_t$	deprec.= $0{,}1 \cdot K_t$	I_N	I (gross)	C_t	Y_t	$Y_t - G_t$
–2	100	1000	1000	100	0	100	800	1000	900
–1	100	1000	1000	100	0	100	800	1000	900
0	102	1000	1000	100	0	100	800	1002	900
1	104.04	1002	1000	100	2	102	800	1006.04	902
2	106.12	1006.04	1002	100.2	4.04	104.06	801.78	1011.96	905.84
3	108.24	1011.96	1006.04	100.60	5.92	106.52	805.19	1019.95	911.71
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
15	137.28	1231.15	1203.97	120.40	27.18	147.58	974.72	1259.58	1122.30
16	140.02	1259.58	1231.15	123.12	28.43	151.55	997.60	1289.17	1149.15

The growth of public expenditure (with a balanced budget) stimulates a growth of Y which induces a growth of I at a rate which grows up to about 2.8% (in periods 11 to 14) and then tends slowly (with damped oscillations) toward 2%, with an associated growth of Y and of K which also becomes initially higher than 2% and then tends toward 2% (it is about 2.3% in period 16); thus after 15 periods the capital stock has grown by over 20%.[5]

Very importantly, the example shows that in order to achieve this result *there is no need for decreases either of consumption, or of the average propensity to consume*: the increase of the rate of growth of capital without decrease of the average propensity to consume is made possible by the increase of Y/K, which makes it possible to increase I/K while also increasing C.

The example shows that there is little reason why this growth process should run against bottlenecks. The average Y/K ratio remains all the time close to 1, arriving at most at about 1.05, which means on average a capacity utilization only 5% greater, something easily obtainable in most situations.[6] The utilization rate *in the capital-goods industry* becomes initially greater than that, of course (because when gross investment increases, it increases initially percentwise more than Y); but a more detailed example, separating sectors and including hypotheses on the forces affecting the allocation of investment among sectors, would have been necessary in order to estimate by how much. What one can say on the basis of this example is that, before the growth process begins, the capital goods industry (assuming there too the capital–output ratio to be 1) employs 1/10 of the capital stock, and that *if* this ratio remained the same afterwards, the utilization ratio in the capital goods industry would rise by 20%; but one can be certain that in fact it would rise much less, because the higher capacity utilization will induce net investment to go in greater proportion toward the capital goods industries, whose productive capacity will therefore increase faster. And anyway, precisely

because of the likelihood of greater fluctuations of demand in the capital goods industry (the accelerator!), this industry can be presumed to be particularly well prepared to adapt production to wide fluctuations of demand.

The reasoning is just as applicable, or even more applicable, to the case where aggregate demand and production were not initially stationary, but were growing at a positive rate g and this rate then becomes g+2%. Even more applicable, because the positive growth rate g justifies the existence of planned underutilised capacity which would not hold in a stationary economy: the construction of bigger plants than necessary at the moment of completion, in the expectation of future increases of demand.

Thus the assumption that, starting from a situation of normal capacity utilization, a faster growth of the autonomous components of aggregate demand (analogous examples might be construed for increases of exports, or of state-controlled investment in nationalized firms) will bring about a 2%-faster growth rate of output and of average productive capacity appears to meet no obstacle in the existing productive capacity.[7] And an increase of two points in the growth rate is quite a considerable increase: already a one-point increase would be considered a great success by most nations.

A simple reversal of the reasoning (assuming that the growth rate of the autonomous components of aggregate demand becomes 2% lower) shows that an insufficient growth rate of demand may be responsible, after a few years, for a very considerable loss of potential productive capacity, a loss easily resulting in structural unemployment, but otherwise not easily perceptible, as it is not visible.

The historical observation that productive capacity is not greatly underutilized for very long periods can then be explained as due to the fact that, if productive capacity is excessive relative to demand, then net investment decreases and may become negative, the older plants are closed down, and productive capacity shrinks or increases at a slower rate than demand, thus tending to adapt to demand.

Except in the event of an *increase* of the growth rate of aggregate demand of more than two percentage points, or of an initial situation where capacity is already utilized more than normally, it seems therefore impossible to agree with Solow on the presumed difficulty demand would have in creating its own supply; the very rapid growth of countries like Korea becomes more easily understandable. For the opposite cases of a prolonged *slowdown* of the growth of aggregate demand, there can be no doubt as to the (cumulative!) effect on the loss of potential capital accumulation.

7.2.4. It might be objected that my example discusses only the adaptation of capital, while – strangely enough – Solow appears to have in mind above all a difficulty with finding additional labour. I say strangely enough, because for an economist like Solow who accepts the neoclassical theory of distribution with the associated conception of capital–labour substitutability, a given level of employment ought not to be an obstacle to an acceleration of capital accumulation. The latter ought only to entail a gradual increase of the ratio K/L.[8]

But even in a non-neoclassical approach, where real wages were considered given and thus also relative prices and technology were essentially given (as implicit in the numerical example presented above), and where therefore a necessity would arise to increase employment if output growth is faster than the growth of labour productivity, Solow's scepticism appears unwarranted in most cases. In the short period, there is nearly universal agreement on the fact that official unemployment rates, besides being always positive, hide the presence of hidden unemployment, and that the rate of participation increases if labour demand increases; this is indeed part of the accepted explanation of Okun's Law; furthermore, employed workers generally do not object to temporary periods of overtime work and wages. If one turns to the ample time intervals relevant for the theory of long-run economic growth, historically one observes clear signs of a (spontaneous or engineered) tendency of the supply of labour to adapt to the demand for labour, so that capitalist economies seem to have always been able to avoid a labour supply constraint. Agricultural underemployment, other pre-capitalist sectors, domestic labour have historically supplied the labour reserves necessary for the industrial revolution and for subsequent growth when population growth was not enough. When that was insufficient (as in post-war Germany, or in the USA), there were huge and carefully regulated immigration flows, or sometimes policies promoting fertility (or, in the opposite case of excessive population, combating fertility, as now in China). It would seem therefore that historically for labour supply too, it is true that in the long run demand has created – spontaneously or through policy interventions – its own supply. The present pressure for immigration from poorer countries toward the industrialized ones suggests that there is little obstacle to the same being true now and in the foreseeable future.

Furthermore the policy question behind all this, i.e. whether the state should or not be assigned a relevant role of regulation of aggregate demand, becomes relevant above all in the periods of recession or crisis, when there is a greater than usual excess of labour supply.

This section was intended to illustrate why long-run output growth may be viewed as determined by the evolution of aggregate demand, with labour supply and capital accumulation *adapting* to aggregate demand, rather than *determining* it via the investment of full-employment savings.[9] I proceed now to show that this different approach to output growth is also relevant for the question of the growth rate of per capita output or labour productivity.

7.3. ENDOGENOUS GROWTH WITHOUT THE FULL EMPLOYMENT OF RESOURCES

7.3.1. If we look at the reasons which in the recent models of endogenous growth make indefinite growth of labour productivity possible even in the absence of exogenous technical progress, we can see that *the force of their action is influenced by the rate of growth of total output; therefore, whether or not this rate is determined by the investment of full-employment savings can make a significant difference to the rate of growth of per capita output.*

One of these reasons can be left aside as unacceptable, namely that proposed by Jones and Manuelli (1990). This is the hypothesis of a CES aggregate production function, which has the property that the marginal product of a factor is bounded above a strictly positive value. It had already been noticed as a possibility by Solow (1956) but then ignored in subsequent literature, evidently because it was judged implausible. Indeed, even neglecting for the sake of argument the illegitimacy of the treatment of capital as a single factor, the hypothesis that the marginal product of capital never goes below a strictly positive value even when labour employment goes to zero appears to have no economically acceptable interpretation.

To summarise, the other reasons used to endogenise the growth of per capita ouput are:

- investment in the production of new knowledge
- investment in the production of human capital
- externalities connected with the expansion of production, or with production itself (e.g. learning by doing)
- division of labour (increase of specialisation) producing scale economies.

Let us then examine the influence of the output growth rate on such causes of per-capita output growth.

7.3.2. Let us start by considering learning-by-doing *à la* Arrow (1962). If one is to leave a role to aggregate demand, then as stated above the full employment of labour should not be assumed (variability in labour employment in the short period is a prerequisite for variability in capacity utilization in response to variations in aggregate demand). Should one then assume a relationship between employment and the real wage? Keynes's approach was to assume a decreasing labour demand curve even in the short period, thus maintaining a univocal decreasing relationship between real wage and Y. But it has been repeatedly noted (e.g. Zenezini, 1990; Brandolini, 1995)[10] that empirical evidence does not accord with this construction for the short period. The Cambridge critique of neoclassical capital theory shows that the labour demand curve is an unacceptable notion even for the long period.[11] On the other hand, the increasingly widely accepted NAIRU approach shows that economists are prepared to admit the role of bargaining strength and other socio-political factors in the determination of the real wage. On this basis, I propose to assume a given real wage (assumed now to increase through time in proportion with labour productivity), which determines technical choices, so that it is as if there were in each period fixed technical coefficients. This has the advantage of allowing us to dispense with a discussion of what should replace that indefensible theoretical construct, the aggregate production function. But in order not to lose all contact with New Growth models, let us still assume a single output (corn produced by corn and labour), as an approximation to a multi-sector economy where aggregation in value terms is possible as long as relative prices are given. Let us then suppose

$$Y = \mathrm{A}K = BL. \tag{1}$$

Let us assume, as is often done, that technical progress is of the purely labour-augmenting type,[12] i.e. that A is not affected by technical progress; Arrow's learning-by-doing (depending, as is well known, on cumulated experience measured by cumulated investment) can then be formalized in this approach by assuming that, at the aggregate level, B is an increasing function of the level reached by the accumulation of capital:

$$B = K^{\beta} \qquad \text{with } \beta \text{ a positive constant.} \tag{2}$$

Hence:[13]

$$Y/L = K^{\beta}. \tag{3}$$

So per capita output increases with the growth of the stock of capital.[14] It is then clear that whatever accelerates the growth of K also accelerates the growth of Y/L. If the growth of capital depends on the growth of aggregate demand which at least partially depends on autonomous elements (e.g. exports, or state expenditure), then the growth rate of aggregate demand will influence the growth rate of Y/L.

7.3.3. Let us now consider the division of labour. Since in the above model it was assumed that Y/K was constant, it would have been equivalent to assume that B was an increasing function of Y rather than of K. This means that the same approach may also be viewed as one possible simple way to formalize a positive influence on labour productivity of the scale of production, and therefore of externalities or increasing returns connected with 'the size of the market'. One sees at a glance – what was intuitive – that a faster growth of output means a faster growth of Y/L.

Romer (1987) tries to capture the increasing-returns effects of an increasing 'division of labour' due to expansion of the size of the market, via a more micro-founded approach. The production function for final output is the same as in the more often quoted Romer (1990): output is produced by labour and by a number of capital goods of different design, whose productivity increases with the increase in the number of designs: the same stock K of capital, if subdivided among a greater number of different types of capital goods, counts as more capital. But rather than the resource–consuming production of new designs as in Romer (1990), it is the scale of aggregate output that determines the number of types of capital goods utilised, owing to U-shaped average cost curves in the production of capital goods, which set a limit to the convenience of specialisation. The functioning of the model is such that in the end it is as if final output were produced by a production function of the usual type $Y = K^{\alpha}(AL)^{1-\alpha}$, where A is the number of different types of capital goods, so one is formally back to labour-augmenting technical progress.[15] So one should find here too that an autonomous role of aggregate demand influences the growth rate of Y/L. Indeed Romer (1987) concludes the analysis of the model as follows: 'any change that leads to an increase in savings – for example a tax subsidy, a decrease in the rate of impatience ρ, or a decrease in the intertemporal substitution parameter σ – will cause growth to speed up; the rate of exogenous technological change will appear to increase' (p. 62). But the increase in savings brings about this consequence only because it increases the growth rate of output. Anything else that increases the growth rate of output will have the same effect.

7.3.4. Let us turn to human capital. In Lucas's formalization of the role of human capital the production of human capital requires no input apart from human capital, and there is no depreciation of human capital, what is not plausible. But these limits of Lucas's formalization are overcome by Rebelo (1991):

$$Y = \psi K^{1-\alpha}(uHL)^{\alpha}$$

$$\dot{H} = (1-\psi)K^{1-\beta}[(1-u)HL]^{\beta} - dH$$

$$\dot{K} = s_K Y - dK$$

where H is aggregate 'human capital'.[16] Rebelo (1991, pp. 508–10) like Lucas assumes that L is constant, and shows that one obtains again Lucas's result, that an increase of the fraction of labour power assigned to the production of human capital increases the growth rate of output and thus, since labour employment is constant, also the growth rate of labour productivity.

Here the role of the full employment assumption is that of making it necessary to choose between more production of current output, and more production of human capital. The admission of the generalized presence of unused resources would mean that it is possible to have more of both, or less of both, depending on the evolution of aggregate demand. The increase in public expenditure considered in the example of Section 7.2.4 might for example be an increase in education expenditure, thus bringing about both a higher growth rate of Y and of H.

But even admitting that the allocation of resources to education is primarily the fruit of private choices, a positive influence of the evolution of aggregate demand on educational attainments in all likelihood exists. The 'production' of skills owing to private investment is responsive both to employment opportunities, and to income, and both are influenced by aggregate demand. The incentive to acquire specializations is higher when there are good job prospects for them; and a higher level of employment is in all likelihood conducive to more schooling, because more families can afford to send their children to school for more years.Also, some training is performed inside firms, and as employment increases, firms will train a greater number of workers.

Therefore a higher level of employment and of Y, and a higher growth rate of aggregate demand, can be trusted to go together with higher average levels of education, and hence, if these are believed to increase labour productivity, with a higher labour productivity.

7.3.5. I turn to investment in the production of knowledge. That the evolution of knowledge may be measured by a scalar is of course nonsense; but even interpreting 'knowledge' as simply an index of labour productivity, and even leaving aside the index problems in measuring the evolution of labour productivity in the face of changes in the quality and kinds of goods produced, still the assumption that increases of labour productivity may be connected by a production function to the resources dedicated to research, and furthermore by a production function which remains unchanged for very long periods, appears at least for the moment to be purely an act of faith which goes against the intrinsic unpredictability of discoveries. Therefore steady growths (or, in more recent literature, constant growths, see e.g. Jones, 2002) derived on the basis of this assumption appear to me to have

very little meaning. Thus even more than for the other arguments I am interested here in the basic insight, rather than in the precise formalizations which have appeared in the Endogenous Growth literature. The basic insight is that advances of knowledge increase labour productivity but require resources to be produced. Here, again, the role of the full employment assumption is to establish a trade-off between producing more output or producing more research; a trade-off which no longer can be presumed to exist if one admits the generalized presence of unused resources whose level of utilization depends on the evolution of aggregate demand. Then more resources dedicated to research can go together with more output: indeed, an increase in output will be the likely effect of more investment in R&D, and conversely a greater output, if the share of output dedicated to research is constant, will mean more research.[17]

In conclusion, on the basis of their own hypotheses as to the determinants of the growth of labour productivity, Endogenous Growth theorists ought to admit that the question, whether one should accept the full (or NAIRU-level) employment of resources, is a central question also for the explanation of the growth rate of labour productivity.

7.4. FULL EMPLOYMENT OR NAIRU? SOME FIRST DOUBTS.

7.4.1. What was argued in Sections 7.2 and 7.3 only indicates the relevance of the question of the determinants of output growth; one might agree, but then add that there are good reasons to accept the full or NAIRU employment of resources in long-run output growth theory.

But is that really so? That this is what a majority of macroeconomists appear to think is obviously no guarantee. Science does not proceed on the basis of majority voting. Let us therefore ask whether the exclusion of a relevant autonomous role of aggregate demand in long-run output growth theory is really solidly based.

Some doubt is immediately raised by the empirical observation of the numerous historical episodes of output growing, for many years in a row, at a rate in all likelihood far lower than that potentially attainable. Besides the Great Depression of the 1930s, one can mention European unemployment since 1974: it is difficult to believe that unemployment rates ten points higher than a few years before entail no potential for greater output expansion. Another obvious case is Japan after 1990.[18]

7.4.2. By themselves, these episodes might be judged insufficient to undermine the theory that market economies spontaneously *tend* to the full employment of resources: it might be argued that this tendency is not very fast, or sometimes meets obstacles which slow it down.

But once this is admitted, then the exclusion of a role of aggregate demand in long-run growth becomes hard to defend: one would have to argue that the periods of underemployment growth are compensated by periods of overemployment growth, so that the trend is sufficiently correctly approximated by an assumption of normal frictional unemployment only (the practical meaning of full employment); and convincing arguments of this type are not easily found.

Let us indeed consider the IS-LM model which, as shown by current textbooks, remains at the core of mainstream short-run macroeconomics. The acceptance of this model makes it possible to argue – if the money supply can be considered sufficiently exogenous – that persistent unemployment is ultimately caused by the downward rigidity of money wages, which impedes the working of the so-called 'Keynes effect' i.e. the effect of variations of the price level on aggregate demand through their effect on the demand for money, hence on the interest rate, hence on investment. The well-known argument is that if money wages decreased in the presence of involuntary unemployment, this would induce a decrease of the price level and then, through the 'Keynes effect', an increase of investment and thus of employment. (The Keynes effect is used against Keynes.) But in order to graduate from this conclusion to the idea that Keynes's analysis is only the explanation of the fluctuations around a trend which is sufficiently close to that of Solow's growth model, it is necessary to argue that money wages are not very rigid, and that the deviations from full employment are not very great, and compensated by deviations of opposite sign; and there is nothing in the model which authorizes such an argument. First, there is nothing in the IS–LM model which authorizes the thesis that, once a period of unemployment caused by a recessionary shock has ended, there will be a period of overfull employment to compensate for it. Second, the observed persistence of levels of unemployment not easily explainable as frictional or voluntary would rather suggest – within this model – that the downward rigidity of money wages is so strong as to maintain the economy far from full employment most of the time. Now, if the level of money wages is given, according to IS–LM theory (as traditionally conceived i.e. coupled, like in Keynes, with a decreasing labour demand curve) the real wage (and hence employment) is determined by the price level which depends on the level of aggregate demand; therefore it is the evolution of the latter which is in the end responsible for the level and rate of growth of output. Furthermore, the working of the 'Keynes effect', which should bring the economy back to full employment if money wages were flexible, relies crucially on the assumption of a significant influence of the interest rate on aggregate investment, and, as

will be argued in Section 7.4, this assumption founders on both empirical and theoretical grounds. But before getting to this issue, it is opportune to discuss the increasingly adopted NAIRU approach.

7.4.3. For theories of growth which deny the relevance of autonomous elements of aggregate demand, the NAIRU approach only implies replacing the full employment of labour with the NAIRU; and it has two advantages. The first is that it does not need a decreasing labour demand curve in order to argue that the economy will gravitate around a definite rate of unemployment; the second is that it is better able to admit a downward stickiness of wages.

But, to start with, the adaptability of labour supply to labour demand (see Section 7.2.4 above) weakens any attempt to determine a NAIRU growth path independent of autonomous aggregate demand influences. The economy of the example in Section 7.2.3 would only need – if there were no labour reserve in the nation – to open up to immigration, and then even accepting the NAIRU theory there would be no inflation, and on the contrary there would be all the advantages for the growth of labour productivity discussed in Section 7.3.

Furthermore, there are empirical and theoretical reasons to doubt the NAIRU approach, which indeed is encountering considerable opposition even among otherwise mainstream economists. Empirically, a growing number of studies concludes that there is no stable relationship between inflation and unemployment, and that the notion of a rate of unemployment, beyond which inflation would continually accelerate, is contradicted by the econometric evidence (see e.g. Setterfield et al., 1992; Rowley, 1995; Galbraith, 1997; Lindbeck and Snower, 1999; Ball, 1999; Solow, 2000). On the more theoretical side, not only doubts have been advanced on the internal consistency and on the stability of the models used to determine the NAIRU (Sawyer, 1997); but also, quite radical general criticisms appear possible. The increasing adoption of the NAIRU approach in place of the monetarist theory of the tendency to a natural rate of unemployment indicates that economists find it increasingly difficult to explain unemployment (in excess of frictional unemployment) as essentially voluntary (in the sense that the unemployed workers are supposedly unwilling to work at the ruling wage), and increasingly admit the need to interpret inflation as mostly cost, rather than demand, inflation and therefore reflecting a distributive conflict. There is thus an interesting *rapprochement* between the positions of mainstream and of conflict economists: unemployment is increasingly viewed as influencing the bargaining strength of wage labour vis-à-vis firms.

But then the undefinability or impermanence of the NAIRU should not come as a surprise, for at least two reasons. First, once it is admitted that firms are compelled to fix cost-covering prices and that therefore, *given the other costs*, an increase in monetary wages greater than the increase of labour productivity obliges firms to raise prices, the 'given-other-costs' clause is clearly crucial, and most of the times illegitimate. Among the other costs there are: interest rates; the salaries of white collar workers, of managers, of external consultants (e.g. lawyers); taxes; public utilities; the prices of imported inputs which depend *inter alia* on the exchange rate. Therefore there are several degrees of freedom which make it possible that there may be no need for firms to increase prices when money wages increase, or contrariwise that firms may have to raise prices in spite of no money wage increases. (That prices increase when employment increases may then sometimes be due to the fact that, fearing inflation, the central bank raises the interest rate!)

The second strictly connected reason is that in the relative bargaining strength of labour and firms many institutional and political elements play a part, such as the preferences of government and of the monetary authorities, the degree of unity of labour and of firms in the bargaining process, the expectations of trade unions on the effects of their action on employment. Thus it is conceivable that trade unions may agree to restrain their demands for wage increases in exchange for policies promoting employment (this is generally agreed to have happened in the so-called 'neo-corporatist' economies). But it is also conceivable that, without considerable variations in unemployment, a change in the political climate may cause sharp increases of wage demands (as in May 1968 in France, or the so-called Hot Autumn of 1969 in Italy) – or decreases, if e.g. there is a change of government in a direction hostile to the labour movement (Pinochet's Chile). So many are the historically-specific elements entering the picture, that it would be indeed surprising if great regularity were to be observed in the connection between unemployment and inflation.

7.4.5. The theory of a spontaneous tendency towards the NAIRU again requires the so-called 'Keynes effect': a level of output higher than the NAIRU level causes – it is argued – an acceleration of inflation (due to money wage increases) which increases the demand for money and the rate of interest, so investment decreases, and unemployment returns to the NAIRU level. Without the spontaneous tendency towards the NAIRU which the Keynes effect should ensure, the thesis that the long-run growth trend is determined by the NAIRU becomes even less credible, because it would require *government intervention* to ensure the tendency to reach the NAIRU, and this is hardly credible, in view e.g. of the long historical periods when concern with unemployment as a regulator of inflation was not found among policy-makers.

Therefore, if the 'Keynes effect' is judged implausible, the mainstream approach to long-run output growth loses all residual credibility, because it loses the mechanism which should ensure Say's Law i.e. the adaptation of investment to the level of savings deriving from the income associated with the full or NAIRU level of resource utilization.[19]

Let us then turn to a discussion of the plausibility of the 'Keynes effect'.

7.5. AGAINST SAY'S LAW

7.5.1. The 'Keynes effect' requires: (*a*) that changes in money wages influence the price level in the same direction; (*b*) that changes in the volume of monetary transactions change the demand for money; (*c*) that changes in the demand for money alter the rate of interest; (*d*) that changes in the rate of interest alter aggregate investment in the opposite direction.

I have argued that (a) is highly debatable (§3.3). Doubts have been advanced on (c) on the basis of 'endogenous money' arguments of various types.[20] Here I discuss (d).

7.5.2. Why investment should be a negatively elastic function of the interest rate is something on which there is at present considerable disagreement; increasing numbers of economists are sceptical about the whole idea.

It is well known that the *empirical* evidence is unable convincingly to support the thesis that the rate of interest exerts a significant influence on investment. This conclusion of older empirical inquiries (see e.g. Junankar, 1972) has not been disproved by later econometric research. The recent survey of investment theory in the *Journal of Economic Literature* concludes:

> While there is clearly no uniformity in the results and the role of shocks remains to be assessed, it appears to this author that, on balance, the response of investment to price variables tends to be small and unimportant relative to quantity variables (Chirinko, 1993, p. 1906; also see Ibidem, pp. 1881, 1883, 1897, 1899).

Thus the empirical evidence would appear to suggest that the negative influence of the rate of interest on aggregate investment, if it exists at all, is too weak to justify the belief that investment adjusts to savings faster than does savings to investment via the Keynesian mechanism of variations of aggregate income. Edmond Malinvaud (1995, pp. 125–7) comes to the same conclusion.[21]

It is only natural, then, to suspect that there might be something unconvincing in the *theoretical* arguments which support the expectation that investment *ought* to exhibit a negative elasticity vis-à-vis the interest rate.

The suspicion proves to be confirmed by theoretical analysis. For reasons of space I can barely mention the problems; for detailed analyses the interested reader is referred to Ackley (1978, chs 18 and 19), and to Petri (1997).

7.5.3. The traditional support for the theory that investment is a decreasing function of the interest rate was the belief in capital–labour substitution, with capital treated as a homogeneous factor, an amount of value. As explained by Garegnani (1990), the traditional demand-for-capital function was in fact the demand for a capital/labour ratio in new plants, and it therefore implied a demand for the (gross) investment necessary to achieve the desired capital/labour ratio in the employment of the flow of labour 'freed' by the closure of the plants reaching the end of their economic life.

This derivation of the investment function was undermined by the discovery in the 1960s of the possibility of *reverse capital deepening* (Garegnani, 1990), which showed that there is no guarantee that the demand for capital is a regularly decreasing function of the rate of interest. But even before this discovery, the theory of investment was in disarray, because the traditional derivation needed a well-defined flow of 'freed' labour, i.e. needed the full employment of labour, which on the contrary could no longer be assumed after the acceptance of the IS–LM foundation of macroeconomics. Hence a number of attempts to derive the decreasing investment function without assuming full employment, all of which suffer from grave deficiencies.

One approach, named 'array-of-opportunities' by Gardner Ackley (1978), argues that entrepreneurs have at each moment in front of them a series of investment projects, which they rank in order of decreasing internal rate of return; they then adopt the projects with a rate of return not lower than the rate of interest; so a lower rate of interest means the adoption of more projects, hence a greater aggregate investment. The basic weakness of this approach is that it treats the returns from the investment projects as given independently of the rate of interest, as if prices could be treated as given. On the contrary, competitive prices tend to equal costs of production; so if the rate of interest decreases, prices will tend to decrease too (relative to money wages) because interest is one of the costs, so the rates of return will tend to decrease as well. Furthermore, competition and the tendency of investment to be allocated in greater proportion to the sectors where the rate of return is greater will tend to annul the differences in rates of return, making all of them *equal* to the rate of interest (plus a risk allowance). So Ackley harshly criticises the 'array-of-opportunities' approach and concludes that the reason why a decrease in the interest rate increases investment can only lie in the induced increase in the desired capital-labour ratio (1978, p. 625, footnote 15). However one then runs into the problems mentioned above, connected

with reverse capital deepening and with the absence of the right to assume the full employment of labour.

Another derivation of a decreasing investment function, making no reference to the 'capital' intensity of production, relies on Kalecki's 'principle of increasing risk' (and more recently has come to be called the liquidity approach). Kalecki (1937) takes product prices as given and argues that the level of investment is limited by the rising 'cost of borrowing' inclusive of risk, because more investment means more borrowing, hence a higher leverage, hence a higher risk of default and bankruptcy. This approach is often used to argue a dependence of investment on sales or cash flow i.e. on demand, but here we are interested in the implication (not much stressed – but admitted – by Kalecki) that a decrease in the basic rate of interest will increase retained profits, and thus induce firms to adopt a higher debt/asset ratio, i.e. to borrow more and invest more. This approach, again, forgets that product prices decrease relative to money wages if the rate of interest decreases, and that therefore a decrease of the rate of interest also decreases the rate of return on investment; retained profits do not increase.

Other approaches which have enjoyed or enjoy considerable popularity are Jorgenson's approach (Jorgenson, 1967) and the adjustment costs approach (see Söderstrom (1976), Abel (1990) for surveys of the vast literature), both of which assume a given number of firms – an incredible assumption in aggregate investment theory! – and also make again the mistake of treating prices as independent of the interest rate, and are therefore clearly unacceptable.

Nowadays Tobin's *q* enjoys greater popularity; but the derivation, from this approach, of a negative dependence of aggregate investment on the rate of interest rests either on adjustment costs, what has been argued to be unacceptable, or on the increasing-supply-price-of-capital-goods approach of Lerner (1944), which is empirically more than doubtful, and which furthermore, in order to explain why a lower interest rate makes the aggregate of firms desire an increase of the capital stock, needs the traditional notion of capital–labour substitution undermined by reverse capital deepening.

In conclusion, the justifications of the view of aggregate investment as a decreasing function of the interest rate either ultimately rely on the indefensible traditional marginalist conception of a decreasing demand function for capital the value factor, or are theoretically indefensible even aside from the criticisms of marginalist/neoclassical capital theory. In particular, again and again the mistake recurs, of treating the yields from investment projects as independent of the level of the interest rate, a mistake pointed out also by the more attentive mainstream theorists, e.g. by Ackley (1978) and even by Jorgenson (1967, p. 152, quoting Alchian, 1955).

7.5.4. There is therefore no convincing theoretical basis for viewing the interest rate as the price which brings investment into equality with savings. Say's Law – the thesis that investment adapts to savings – loses its foundation. The faith in the spontaneous tendency of market economies towards the full (or NAIRU) employment of resources must therefore be abandoned.

It is then natural to turn again to the principle of effective demand, i.e. to the idea that it is variations in income which bring about equality between savings and investment or more generally between aggregate income and aggregate demand. Output and its growth must then be explained through the evolution of the autonomous components of aggregate demand. This will in all likelihood make it easier to reconcile economic theory with historical episodes such as the Great Crisis of the 1930s, or persisting European unemployment, or the East Asian initial successes and present difficulties. The theory of *per capita* growth will have to recast its insights accordingly.

NOTES

* Funding from the Ministero dell'Università is gratefully acknowledged. I thank Mauro Caminati and Carlo Panico for their comments.

1. This explains why little attention is given to distinguishing per capita output from labour productivity i.e. output per unit of employed labour. (This is much less acceptable when one admits the possibility of ample variations of unemployment rates, and the possibility of an influence of demand on participation rates, but these issues will not be discussed here, and for brevity's sake the term 'per capita output' will be used below to indicate average labour productivity).
2. Say's Law states that investment adapts to savings rather than vice versa; as made clear by e.g. Ricardo, it does not imply the full employment of labour; it does so only in the neoclassical approach to distribution, which postulates a decreasing demand curve for labour and where Say's Law guarantees that there will not be aggregate demand deficiencies to create obstacles to increases in employment when, owing to unemployment, real wages decrease and the demand for labour therefore increases. On the other hand, the Pigou or real-balance effect might bring about the full employment of labour without the validity of Say's Law. This is why it was found necessary in the title to cite both Say's Law and the full employment of resources.
3. Jacob Steindl has particularly insisted on the first group of reasons, Robin Marris on the second. Another possible reason for excess capacity in oligopolistic industries is entry deterrence.
4. On the contrary the usual simplification is maintained of giving no role to inventories or to their fluctuations.
5. What happens is that Y and K grow at a higher rate than G until the G/Y ratio has sufficiently decreased so as to make room for the higher I/Y ratio necessary for a 2% rate of growth with normal capacity utilization. In this example the multiplier–accelerator interaction does not cause instability.

6. U.S. Census Bureau data on capacity utilization for the fourth quarters of 1997 to 1999 (http://www.census.gov/prod/2001pubs/mqc1-99.pdf) show an average utilization rate for all manufacturing of respectively 76%, 73% and 74%, which can be compared with an average utilization rate over the period 1967–99 of about 82% (and 'popular wisdom' that inflation starts when utilization rates go above 85%). Presumably little difficulty would have been found in those quarters with producing even 10% more; the constraint was clearly on the demand side. These are measures of the utilization rate relative to 'normal' production, but firms also indicate a 'national emergency' potential production from 40% to 80% higher than this normal production.
7. Other obstacles might be a labour supply constraint (discussed below) or an inflation constraint (see Section 3 below); a balance-of-payments constraint would on the contrary be proof of the influence of aggregate demand on growth.
8. And for the short run, it is Solow himself who has repeatedly argued that the observed unemployment is largely involuntary, implying little difficulty with increases of employment.
9. See Garegnani and Palumbo (1998) for further implications of such an approach and for some remarks on historical evidence supporting it.
10. Among older studies, Dunlop's and Tarshis' objections to Keynes are well known.
11. The Marshallian long period is that situation in which, owing to the greater speed with which the relative quantities of different capital goods can change relative to the speed of accumulation, one can treat the composition of capital as endogenously determined, while neglecting the effects of accumulation. But then, how should one specify the given employment of capital? The given capital stock is indeterminable, because since its composition must be treated as variable, it can only be measured as an amount of value; and a given value of capital K is an indefensible notion because values change with the real wage. So it is impossible to determine the labour demand curve.
12. This obviously unrealistic hypothesis, like others in this section, is only made for the purpose of departing as little as possible from Endogenous Growth models.
13. Romer (1986) assumes a knowledge-producing activity of firms, with knowledge spillovers to other firms, but the final result is, as is well known, extremely similar to Arrow's and, if one assumes fixed coefficients as here in equation (1), then equations (2) and (3) appear also to represent Romer's ideas as to the final result of what in that article he sees as the causes of technical progress (with the difference that Arrow would assume $\beta < 1$ while Romer would assume $\beta > 1$). This may be the place for a comment on Romer's article. Romer assumes that the consumption good is produced (by a given number of firms) by firm-specific knowledge k, general knowledge K, and a vector x of other inputs given in amounts; and that each firm produces private knowledge via the use of its own private knowledge and of forgone consumption. The details of these assumptions are difficult to justify, and yet no justification is offered: for example, why doesn't general knowledge K also influence the production of private knowledge? Why aren't the rentals of the given factors included in the costs? Why is there no mention of other factors in the production of knowledge? Why (in a model of long-run growth!) is the number of firms given? Can we measure knowledge quantitatively so as to be able to say that knowledge has e.g. doubled? Should an article be accepted for publication when assumptions as dubious as these receive no justification?
14. If the growth rate of K is constant, then the growth rate of Y/L is constant, decreasing, or increasing depending on whether β is equal to, less, or more than 1, but this is of limited interest here; even when $\beta < 1$ there is indefinite growth of Y/L, and the reason for wanting to obtain a constant growth of Y/L is only the determination of steady states, whose empirical relevance is far from established and a priori highly doubtful on theoretical

grounds, given the difficulties of aggregation and of intertemporal comparisons of a *Y* including all the time new products.

15. '... the economy will behave as if there is a form of exogenous, labour-augmenting technological change' (Romer, 1987, p. 61–2).
16. It is striking how little space is devoted in the New Growth literature to justifying crucial assumptions. How legitimate is the representation of human capital as influencing the efficiency units of labour through a multiplicative effect HL? Neither Lucas nor Rebelo pose the question. Treating human capital as a factor of production in the usual sense would imply that human capital should run against decreasing returns when added to a constant quantity of labour. The multiplicative approach has been probably aided by the inherent vagueness of the notion of 'human capital', which has to do with acquiring know-how, something different from increasing the amount of an input. The correct analogy seems to be with software: adding human capital is similar to adding more or better software to a computer. But then different 'quantities' of human capital mean that one is dealing with different kinds of labour, and aggregability and measurement of increases of the stock of human capital are highly dubious. Several authors write that this multiplicative approach has been suggested by studies (e.g. Mincer) showing that for each extra year of schooling wages increase by approximately the same percentage, 10% for some authors, 7% for others. A non-neoclassical economist would argue that wage differentials are not due to differences in marginal products, and the capital-theoretic criticisms support such a stance. But even within a neoclassical framework that justification appears questionable. Education changes one's skills, i.e. changes the kind of labour one offers, so in a supply-and-demand approach relative wages depend on the relative scarcities of the different kinds of skills, and it will be the choices of individuals which will alter relative scarcities until an equal convenience of investment in education is reached; in other words, the 'marginal product' of an extra year of education might well be decreasing in the sense that an increase of *K* per person and of the 'amount of education' (assuming it were measurable) per person in the same proportion might increase output less than proportionally, but in an economy where different degrees of education coexist, since different skills are not perfectly substitutable, if the supply of the more highly skilled labour is sufficiently decreased, its marginal product and wage will increase, and this will indeed happen until the rate of return on one more year of education becomes the one desired by consumers: so within a neoclassical growth framework the observed 'marginal product' of education reflects consumer choices and not a technological property of that 'marginal product'.
17. The 'non-scale' recent literature (see Jones, 1999, for a simple characterization) argues that, given the share of output going to research, a bigger output does not mean a higher rate of growth of labour productivity if the increase in output goes together with a proportional increase in the variety of goods produced, and if research is sector-specific and with no spillover to other sectors. No doubt the reasoning behind this argument has some relevance, but it can hardly be used to deny that a larger economy dedicating more resources to research will be advantaged by its greater size. The different size of the economy is what made it easy for the USA to get ahead of the USSR in military technology, by allocating to military research a much larger budget. Also, a bigger economy does not generally have a proportionally greater number of no-spillover sectors; e.g. in the USA many industries are simply bigger than in European nations.
18. I am not claiming that mainstream macroeconomics is unable to find explanations for these episodes; I only claim that resources were underutilized and therefore policies, maintaining aggregate demand higher than it was, would have ensured a higher utilization of resources and higher growth rate.

19. There is general consensus (see e.g. Patinkin, 1987) that the real-balance effect is too weak, and too fraught with difficulties (e.g. bankruptcies) caused by the required price decreases, to be relied upon as a plausible mechanism ensuring a sufficiently fast tendency of aggregate demand to increase when prices decrease. A recent calculation by Sawyer (1997, section 6) based on the NAIRU model of Layard, Jackman and Nickell concludes that a decrease of the fiscal stance causing a 1% decrease of aggregate demand would require a 67% decrease of money prices in order for the real balance effect to counterbalance it – assuming no perverse effects of the price decrease on investment.
20. Besides the well-known modern defenders (e.g. Kaldor, Basil Moore) of the endogeneity of the money supply owing to overdraft facilities or to the creation of money substitutes, it is interesting to remember that recently David Romer (2000) proposed to give up the LM curve of the IS–LM model owing to the observation that central banks increasingly target the interest rate rather than the money supply. Pivetti (1991, ch. 2) summarises evidence suggesting that this is not only a recent tendency as Romer suggests, but a nearly universal aspect of capitalism in industrialised countries.
21. See also Hall (1993, pp. 278–9): 'established models are unhelpful in understanding this [1990–91] recession, and probably most of its predecessors. ... In spite of low interest rates, firms cut all forms of investment. ... Little of this falls into the type of behaviour predicted by neoclassical models'.

8. The demographic transition and neo-classical models of balanced growth

Piero Manfredi and Luciano Fanti

8.1. INTRODUCTION

An aspect lying at the very heart of classical economics, which is lost in neo-classical theory, is, as Samuelson (1985) sharply claimed, the endogenous interaction between population and the economy,

> Once upon a time, throughout the heyday of classical economics, demography belonged to political economy. The supply of labour was one of the important endogenous variables in the systems of Smith, Malthus, Mill and Marx. One feature of neoclassical economics that distinguishes it from the classical version, is the removal of population as a variable subject to economists' analysis (p. 166).

This 'removal' is evident in the 'standard' neoclassical model of economic growth (Solow, 1956), in which the production factors, labour and capital, are fully exogenous. Such a switch in the focus of analysis, the roots of which date back to the birth of neoclassical theory, is only partly a consequence of the 'logical structure' of neoclassical theory (Kurz and Salvadori, this volume, ch. 1). In fact already in his 1956 paper Solow stated that he was dissatisfied with the gross treatment given to labour, and suggested the need for fully endogenising the dynamics of population and the supply of labour. An important, though obvious, consequence of population endogenisation is that Solow's model (and in general all 'proper' neo-classical models) becomes, according to the taxonomy in Kurz and Salvadori (this volume, ch. 1), an endogenous growth model.

In this paper we discuss some of the effects of endogenous population dynamics and labour supply within Solow's model. We will do this by using the 'case study' of a topic which is at the core of endogenous growth theories with population, namely the economic explanation of the Demographic Transition (DT). The Demographic Transition currently represents a major challenge for scholars in demo-economics aiming to find endogenous explanations of the transitions between great historical regimes (see Galor

and Weil, 1999). A fairly accepted view is that Solow's (1956) model is hardly able to provide any explanation of the transition. This is expressed, for instance, by Galor and Weil (1999), and Chu (1998, p. 133). The latter, emphasising the gain permitted by 'new growth' models, such as the Becker–Murphy–Tamura model (Becker et al., 1990), states:

> ... the explanation of the Demographic Transition has not been successful under the neo-classical model of Solow, for it typically predicts a converging steady-state growth rate of per capita income, which is incompatible with the diverging development paths among countries observed over the past 50 years. Moreover, Solow's model is also weak in predicting the relationship between income growth rate and population growth rate. It is well known that in Solow's model the steady state rate of per-capita income is a decreasing function of the population growth rate.

Though generally accepted, the above reasoning is, to our mind, exaggerated. What chiefly matters is not whether Solow's original model with its exogenous supply of labour can predict the demographic transition, or explain its timing and forms, obviously doesn't. Rather the point is whether, adequately equipped, i.e. by postulating sound 'transitional' hypotheses, Solow's model can offer insights into fundamental aspects of modern growth.

The present paper is divided into three main parts. In the first part we review some main aspects of the DT aiming to combine the distinct perspectives of historical demography and growth theorists. In the second part we critically discuss the main strategies used to incorporate the Transition in economic models. Here the DT is essentially regarded as a major change in mortality and fertility patterns, leading to a peculiar time profile in the relationship between the growth rate of the population and output per-capita.[1] We distinguish two main approaches: 1) the transition is microfounded, as the outcome of the maximising behaviour of individuals (for instance Jones, 1999); 2) the transition is postulated on empirical bases, emphasizing the asynchronous patterns of decay of mortality and fertility as per-capita income increases (as in Strulik, 1997, 1999a, 1999b). We also use a 'diffusionist' argument to model the asynchronous decay of mortality and fertility. The 'diffusionist' approach (Rosero-Bixby and Casterline, 1993) has often proved to be more effective than demand-supply mechanisms in explaining the transition.

Finally, in the third part we investigate the effects of transitional assumptions on simple neo-classical models in the presence of both constant (CRS), and decreasing (DRS) returns to scale. First we consider a traditional CRS Solow model. In the absence of technical progress, mere population mechanisms lead to the existence of multiple equilibria, with a stable 'poor equilibrium', i.e. a poverty trap, coexisting with a stable rich ('modern') equilibrium. Moreover, the well-known ‘absolute convergence’ statement of Solow’s model is replaced by the more general result by which countries that escaped the poverty trap with different endowments of capital would possibly experience an initial divergent phase, which should convert to a convergent regime only in the long (perhaps very long) run. This result provides a powerful explanatory tool for observed dynamics, and an answer to the critical difficulty of Solow’s model previously quoted from Chu. We subsequently consider a basic neoclassical model with DRS. DRS lead, compared to CRS, to important, and more empirically plausible, additional results, which have not, with the exception of Strulik (1999a), been stressed in the literature on growth: a) coexistence of stable poor ‘Malthusian stagnation’ equilibria with stable states of ‘modern’ long term balanced growth, as historically observed, suggesting that DRS models are better than CRS as candidates for unified modelling of ‘very-long’ term economic growth; b) appearance of persistent endogenous oscillations around the ‘Malthusian stagnation’ equilibrium when the population transition is considered jointly with the, quite well documented, saving transition. The latter result, obtained via the most standard growth model, offers a pleasing alternative perspective, i.e. a purely endogenous one, to the traditional exogenous explanations of oscillations around the stagnant Malthusian equilibrium. It is also interesting for the debate on Malthusian oscillations, an important topic in the recent demoeconomic debate (Lee, 1997).

This paper is organised as follows. The DT is reviewed in Section 8.2. Section 8.3 discusses some different modelling frameworks of the transition. Sections 8.4 and 8.5 study some CRS and DRS Solow-type models with DT, and are followed by concluding remarks.

8.2. THE DEMOGRAPHIC TRANSITION AND ITS ECONOMIC DETERMINANTS: A REVIEW

By the phrase Demographic Transition demographers mean the set of dramatic changes by which, starting in Western Europe, particularly England and France, during the 18th century, demographic systems moved from their Ancien Régime, in which stationarity was maintained via a highly ‘expensive’ balance between high levels of fertility and mortality, to their modern regime, characterised by low fertility and mortality (Chesnais, 1992; Livi-Bacci, 1997, and references therein). Following the standpoint of historical demography (Livi-Bacci, 1997, 1998) the major engine of the whole process appears to be the well-documented achievement of a threshold in the level of technology (or an increase in its pace of growth) which, allowing an increase in agricultural surpluses, led to: *a*) a sharp ‘endogenous’

reduction in mortality; *b*) an increasing trend in per-capita income and wellbeing, *c*) a fall in the demand for labour in agriculture which was the condition for a transformation of social organisation, from rural to urban. The reduction in mortality favours the subsequent reduction in fertility,[2] via the great shift from quantity to quality of children (Galor and Weil, 2000). The diffusion of birth control was the main tool through which fertility decline became possible. The outcome of this process was a transient period of fast population growth, due to the asynchronous decay of mortality and fertility, and per-capita income. Population growth eventually came to an end, due to the continued decrease in fertility, whereas growth in per-capita income has been uninterrupted.

Demographic approaches are mainly descriptive, focusing on heterogeneities in timing at onset, pace, and local features of the transition, and no attempt has been made to develop simple formal demoeconomic explanations of the transition. Major questions concerning, for example, the inevitability of the transition, have not been posed at all. Very recently the transition has become the object of renewed interest by (demo-) economic growth theorists (Galor and Weil, 2000 and refs therein; Jones, 1999 and refs therein).

In order to manage what has become a jungle of papers with a huge number of different definitions, we introduce some fundamental definitions into the endogenous growth theory with population, e.g. with respect to the terms of ‘poverty trap’ and ‘demoeconomic regime’.

‘Poverty trap’ traditionally (Barro and Sala-i-Martin, 1995) refers to a stable steady state whose ‘name’ can be motivated thus: 1) ‘poverty’ because it has low levels of per capita output and capital stock, 2) ‘trap’ because if agents attempt to break out of it then the economy always tends to return to it. But the meaning of ‘trap’ is clearer if it co-exists with (at least) another equilibrium with better properties in terms of ‘welfare’ which is either i) repelling (so that when the economy could – due to some shock – approach it, the trap would inexorably re-swallow it), or ii) attracting but with a basin of attraction so far from the ‘poverty trap’ region as to be unattainable for the economic variables. The search for models with endogenous mechanisms of escape from the poverty trap has been a major ‘target’ for growth theorists.[3] It is easy to see that, by only focusing on the dynamics of per capita physical capital as in the Solowian model, in order to attain the ‘target’ (for instance in order to have the interval of rising average product, necessary to have multiple equilibria) there are few possibilities: i) of assuming increasing returns to scale for all the factors through learning by doing or other externalities, 2) of assuming a non-linear saving function, 3) of suitably endogenising population growth. While the first two candidates have been widely used, and in particular the first perhaps is, directly or indirectly, the 'growth engine' of all the endogenous growth theory, the third, however widely explored, has not been successful, to our knowledge, within the simple framework of Solow's model (1956). In the present work we attempt to fill this gap.

According to Galor and Weil (1999) three distinct regimes have characterised the process of economic development: the 'Malthusian' Regime, the 'Post-Malthusian' Regime, and the 'Modern Growth' Regime. To fully understand these definitions it is useful to distinguish the macroeconomic point of view from the demo-economic one: the first focuses on the behaviour of income per capita and technological progress, while the second on the relationship between the level of income per capita and the growth rate of the population.

- *Malthusian Regime*:
  1) low technological progress and population growth, at least relative to modern standards, and roughly constant income per capita;
  2) *positive* relationship between income per capita and population growth.
- *Post-Malthusian Regime*:
  1) growing income per capita during this period, although not as rapidly as during the Modern Growth regime;
  2) positive relationship between income per capita and population growth (still as in the Malthusian Regime).
- *Modern Growth Regime*:
  1) steady growth in both income per capita and the level of technology;
  2) negative relationship between the level of output and the growth rate of population.

From the demoeconomic point of view the DT is represented by the transition from the Malthusian to the Modern Growth Regime. As we are mainly interested in the relation between economic growth and DT, in this section we essentially consider the two latter regimes. The Post Malthusian Regime would be interesting if the focus were on the Industrial Revolution, technological progress and other macroeconomic aspects.[4]

Obviously research on the interaction between income growth and fertility is not new in economics (e.g. Razin and BenZion, 1975; Cigno, 1981, 1984a, 1984b; Becker, 1988) but only more recently the literature has focused on the existence of different long-run regimes corresponding to major epochs of mankind's history[5] (amongst others Becker et al., 1990; Azariadis and Drazen, 1991; Ehrlich-Lui, 1991; Kremer, 1993; Palivos, 1995; Tamura, 1996; Yip and Zhang, 1997; Lucas, 1998; Strulik, 1997, 1999a; Dahan and

Tsiddon, 1998; Galor and Weil, 1996, 1998, 1999; Jones, 1998, 1999; Hansen and Prescott, 1999; Kögel and Prskawetz, 2001).

8.3. REPRESENTATIONS OF THE DT IN GROWTH MODELS: OPTIMISING VERSUS DIFFUSIONIST PERSPECTIVES

The present section discusses some main avenues through which the DT may be modelled: 1) via full microfoundation; 2) via an empirically-based formulation.

Moreover, the empirically-based avenue can emphasize either only the asynchronous patterns of decay of mortality and fertility as per-capita income increases or the more general 'diffusionist' mechanism.

As we will see, in the final analysis, all these approaches lead to endogenising the supply of labour via a humped function of per-capita income, that will be used for subsequent analyses in descriptive growth models à la Solow.

8.3.1. The Demographic Transition as the Outcome of an Optimal Choice

Here we enrich Jones' (1999) formulation by adding an exogenous non-zero saving rate, in order to make Jones' formulation compatible with the accumulation side in Solow's (1956) model. Each individual has, in each period of time, an endowment of one unit of labour which can be used to obtain consumption or children. Let h denote the fraction of time spent working ($1-h$ = time spent producing children), w the earnings per unit of time worked, s the (constant) saving rate, $c°$ the (constant) subsistence consumption, $b°$ the constant number of children independent of the time spent on childrearing (i.e. a 'natural' fertility). According to Jones (1999) the utility of individuals depends on consumption (c) and number of children (b) as follows

$$u(c,b)=\frac{(1-m)(c-c^{\circ})^{1-\gamma}}{1-\gamma}+\frac{m(b-b^{\circ})^{1-\eta}}{1-\eta} \qquad 0<m<1 \quad ; \quad 0<\gamma,\eta\leq 1 \qquad (1)$$

The individuals take w as given and solve the problem:

$$\max_{c,b,h} u(c-c^{\circ},b-b^{\circ}) \qquad (2)$$

subject to the restraints:

i) $c=(1-s)wh$; and

ii) $b=f(1-h)$;

The restraint *ii*) states that each unit of time spent producing children produces f births, with $f > b°$.

A simple reformulation of constraints *i*) and *ii*) gives the usual budget constraint

$$c+\frac{b(1-s)w}{f}=(1-s)w \tag{3}$$

where $(1-s)w$ is disposable income and $1/f$ is the (per child) cost of childrearing. Jones has shown that the relation between the optimal fertility rate and the wage, which in general is only defined in implicit terms, actually defines b as a humped function of w, therefore mirroring a major stylised fact of the Demographic Transition (see Jones, 1999 for details). In order to obtain a greater analytical insight, we develop Jones' formulation for the special case[6] $\gamma=\eta$, corresponding to the well-known utility function of the 'Constant Elasticity of Substitution'-type, which allows the explicit solution:

$$b=\frac{b^{\circ}f+Hf\left((1-s)w^{\eta-1/\eta}-c^{\circ}w^{-1/\eta}\right)}{\Omega} \tag{4}$$

where H is a suitable positive constants and Ω is a positive function of the wage.

A differentiation of (4) shows that

$$\frac{\partial b}{\partial w}=\frac{Hf\left[c^{\circ}-w(1-\eta)(1-s)\right]}{\eta w\Omega}+\frac{Hf\left[c^{\circ}-w(1-s)\right](1-s)(1-\eta)w^{-\frac{2}{\eta}}}{\eta\Omega^{2}}+$$
$$+\frac{Hfb^{\circ}(1-s)(1-\eta)w^{-\frac{1}{\eta}}}{\eta\Omega^{2}}$$

The investigation of the latter expression suggests that, *ceteris paribus*, a high subsistence consumption, a high propensity to saving, and a high 'natural' fertility mean that the hump in fertility occurs for larger wage values. This implies that these factors could have played a role in temporally delaying the onset of the population bulge (i.e. the attainment of the moment of maximal population growth). In other words differences in 'cultural' (or 'preference') factors such as the level of perceived minimal well-being,

'natural' fertility and saving behaviour can explain different time-paths as regards the DT in different countries (and ultimately, as we will see in the next sections, also the escape of poverty and self-sustained growth).

The previous development can be used to endogenise population growth in Solow's (1956) model, in which the supply of labour L is assumed to be exogenously growing. In general terms the quantity of labour supplied for production is related to the total population in the work age span (N) by $L=hN$ where h is the participation rate. By taking h as constant over time, it simply holds that $\hat{L}=\hat{N}$. Let us assume (under some simplifying assumptions) that $\hat{N}=b(w)-\mu(w)=n(w)$, where b and μ respectively denote the birth and death rates in the population. Obviously, even in the simplest case, i.e. μ taken as constant, the one-hump shape of b implies that n is one-hump as well. If n is also non-negative, which can be obtained by simply taking $\mu = b_0$, then $n(w)$ faithfully mirrors the demographic transition.

The present approach offers a route to endogenous modelling of fertility transition within descriptive growth models which, being in full closed form, offers potentially interesting analytical results.

8.3.2 Empirical Approaches to the DT: Asynchronous Decay of Fertility (and Mortality) with Per-capita Income, and the 'Diffusionist' Perspective

The simplest 'model' of the DT used by demographers represents the transition as a historical phase of population growth separating the ancient and modern demographic eras, which are mainly characterised by stationarity. This phase of growth is the outcome of the asynchronous decline, usually s-shaped, of mortality (falling first) and fertility over time (Chesnais, 1992). This empirical fact, together with another empirical fact – the monotonically increasing time trend of per-capita income during the same historical period – implies that the corresponding rate of change of the population, i.e. the difference $n(y) = b(y) - \mu(y)$, turns out to be a humped function of per-capita income (as in the case of the microfounded behaviour shown in Section 8.3.1). This humped function proves to fit classical 'transitional' data sets very well (Strulik, 1997; 1999a).[7]

The dynamic consequences of a humped $n(y)$ function are far from trivial. As Strulik (1997) correctly noticed the traditional demographic approaches to the DT are basically a descriptive comment of the peculiar time patterns of the components of population growth observed during the transitional regime. This may lead to the wrong impression that the transition has simply been 'inevitable'. As shown in the next section, even the simplest neo-classical growth model suggests that this does not need to be the case.

Though the choice of s-shaped fertility and mortality functions over time appears mainly motivated by empirical considerations, there is an important theoretical argument in favour of the 'logistic curve' approach of the previous section, namely the diffusionist argument. There is a strong body of evidence (Livi-Bacci, 1997; 1998) suggesting that the main tool by which fertility decline was made possible during the DT was the diffusion of the practice of birth control. The diffusionist paradigm is often taken as the basis for alternative explanations to those based on economic factors, which often provide better fits of observed patterns and pace of the transition (Rosero-Bixby and Casterline, 1993, and references therein).[8] Here we consider a model for the diffusion of birth control which justifies under very wide conditions the s-shaped time pattern for the birth rate during the transition. Consider a population in which the transitional decrease in mortality has already occurred, and which is composed by two groups, those who are still practising natural (uncontrolled) fertility, and those who are practising birth control. We assume that the 'natural' group has high fertility and would, if left alone, grow exponentially, whereas the 'controlled' group has a lower fertility. Let $X = X(t)$ and $Y = Y(t)$ respectively denote the size of the natural and controlled fertility subgroups. Individuals are assumed to move from the natural to the controlled group due to the diffusion of information on birth control, which occurs via both external (the action of the media) and internal (inter-human contacts) diffusion.[9] Moreover, children are assumed to inherit the fertility attitudes of their parents at birth. The model is the following:

$$\begin{cases} \dot{X} = b_X X - \mu X - \left(\alpha X + \beta \dfrac{XY}{N}\right) \\ \dot{Y} = b_Y Y - \mu Y + \alpha X + \beta \dfrac{XY}{N} \end{cases} \tag{5}$$

where μ is the mortality rate; b_X, b_Y are, respectively, the fertility rates of the natural and controlled groups, with $b_X > b_Y \geq \mu$; α is the rate of transition from the natural to the controlled group due to external information, β is the contact rate in internal diffusion. Model (5) is more general than other models used for the diffusion of birth control in that it is not limited to the situation of a stationary population (as in Rosero-Bixby and Casterline, 1993).

What does model (5) predict? Let $r_X = b_X - \mu - \alpha$; $r_Y = b_Y - \mu$ respectively denote the rates of growth of the two subpopulations in the absence of diffusion. We assume $r_X \geq r_Y$ because the opposite case would imply that external diffusion alone is capable of quickly driving the natural group to a slower growth compared to the controlled group. Let us consider

as new variables the total population $N = X + Y$, and the controlled fraction $\varphi = Y/N$ $(0 \leq \varphi \leq 1)$. We obtain the decoupled equations:

$$\dot{\varphi} = (1-\varphi)\left(\alpha + \left(\beta - (b_X - b_Y)\right)\varphi\right) \tag{6a}$$

$$\frac{\dot{N}}{N} = r_Y + (b_X - b_Y)(1-\varphi) \tag{6b}$$

with the initial condition $\varphi(0) = 0$. There are two qualitatively interesting cases.[10] First, if $\beta > b_X - b_Y$, then (6a) only has the non-trivial equilibrium $\varphi_1 = 1$ which is globally asymptotically stable (GAS). In other words, when the effects of inter-human communication are sufficiently strong, then in the long term all the population choose, in relative terms, to move to birth control. This means that the higher fertility of the natural group is made ineffective by the 'migration' toward the lower fertility group. If $r_Y = 0$ the total population becomes stationary in the long term, whereas it continues to grow if $r_Y > 0$. Second, if $\beta < b_X - b_Y$, (6a) also has the equilibrium $\varphi_2 = \alpha/[(b_X - b_Y) - \beta]$ which is meaningful for $\beta + \alpha < b_X - b_Y$, i.e. when the overall rate of information diffusion is below a prescribed threshold. In this case φ_2 is GAS, implying long-term coexistence between the two groups.

Consider now the overall birth rate in our population, which by definition is a weighted average of the birth rates of the natural and controlled fertility subgroups: $b_N(t) = b_X(1-\varphi) + b_Y\varphi$. Note that in both cases considered above, the time behaviour of $\varphi(t)$ will be that of an s-shaped function very much like the logistic curve. In the first case for instance, postulating $\varphi(0) = 0$, the overall fertility rate b_N will decline logistically over time from its natural fertility level b_X to its controlled level b_Y, as largely documented by the empirical evidence on the DT.

This result moreover suggests that the s-shaped relations between fertility rate and per-capita income (rather than time) used for instance by Strulik (1997, 1999a, 2000) necessarily appears empirically due to the monotonically increasing historical time trend of per-capita income during the transitional era.

8.4. THE DEMOGRAPHIC TRANSITION AND SOLOW'S MODEL: THE CASE OF CONSTANT RETURNS TO SCALE

In this section we investigate the dynamic consequences of the DT, within the Solow (1956) model without technical progress. According to the discussion of Section 8.3, the DT may be, as a first approximation, represented by replacing the exogenous rate of change of the supply of labour with a positive humped function $n(y)$ of per-capita income y. The ensuing 'transitional' Solow-type model is

$$\dot{k} = s(y) f(k) - (\delta + n(y))k \tag{7}$$

where k denotes the capital labour ratio, $f(k)$ is a CRS production function (in per capita terms) satisfying Inada's condition, s is the saving rate (here postulated, in contrast with the original Solow's model, to be income-dependent, but will be simply assumed constant below), δ is the rate of capital depreciation, $n(y)$ is the rate of change in the supply of labour. Since $y = f(k)$ under CRS, we obtain (by writing $s(k), n(k)$ as shortcuts for $s(f(k))$, $n(f(k))$:

$$\dot{k} = s(k) f(k) - (\delta + n(k))k \tag{8}$$

Clearly n has a humped relation with per-capita capital as well. We assume that n starts from zero (or a slight positive value) at the beginning of the transition, increases up to a maximum, and finally decreases up to a small positive (or zero) asymptotic value at the end of the DT.[11]

Remark 1. Formulation (8) is well acknowledged in the literature on economic growth with endogenous populations. Already Solow (1956) considered it (see also Nelson, 1956, Niehans, 1963), though he concluded that population endogenizaton, regardless of the complexity of $n(.)$, had no relevant effects on his main results. Nerlove and Raut (1998, p. 1127) emphasize this point, thus raising strong skepticism about the approach:

> it is clear that merely endogeneizing population growth at the macro-level does not shed light on the shape of n and thus on the nature of the dynamics; a utility-maximising model should be used to elucidate the nature of the function.

Their remark is in principle correct and indeed we have fully elucidated the microfounded nature of the shape of n in the previous Section 8.3.1 (see also for other different attempts to give microfoundations to the shape of n, Momota and Futagami, 2000 and Strulik, 1999b).

Nevetheless, we believe that an indisputable exception to the claim of Nerlove and Raut is represented by the humped 'demographic transition' hypothesis which is supported by extensive empirical evidence and thus may also justify the macro-level approach to the DT.

The analysis of (8) is straightforwardly carried out by the usual graphical tools (Barro and Sala-i-Martin, 1995). Let us consider (Figure 8.1) the case

of a constant saving rate. The function $m_1(k) = sf(k)/k$ has the traditional decreasing form whereas $m2(k) = \delta + n(k)$ mirrors the humped form of $n(k)$. The vertical distance $m1(k) - m2(k)$ is the rate of growth of k over time, denoted as γ_k. Figure 8.1 shows that multiple equilibria may exist. The behaviour of (8) is summarised by the following:

Proposition 1. Apart from the zero equilibrium, system (8) admits one or three non-trivial equilibria. In the former case let k_1 be the unique positive equilibrium. Then k_1 is always globally stable. In the latter case (see Figure 8.1) let the three equilibria be $k_{low}<k_{mid}<k_{high}$. It is easy to show that k_{mid} is unstable, whereas k_{low} and k_{high} are locally stable, with respective basins of attraction: $Bas(k_{low}) = (0,k_{med})$, $Bas(k_{large}) = (k_{mid},+\infty)$.

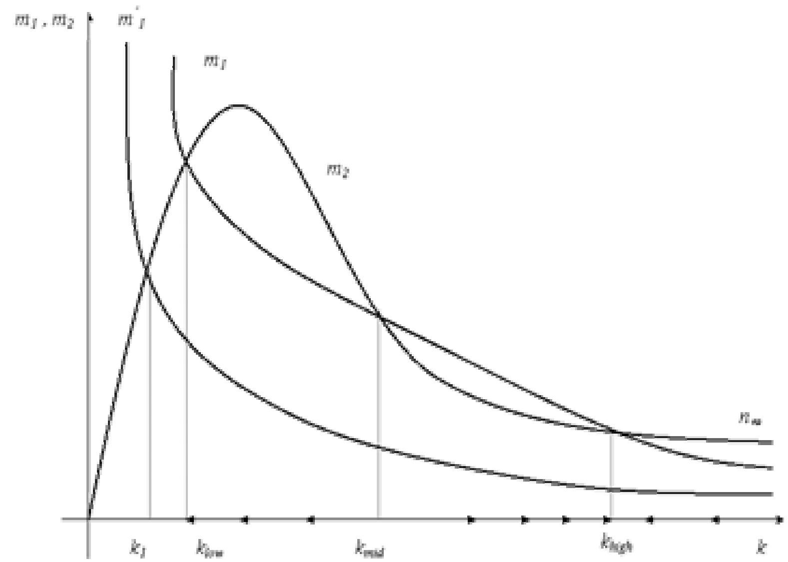


Figure 8.1 – The three equilibria in Solow's model with transitional dynamics

The proof easily follows from examination of the sign of the rate of change γ_k (the arrows on the horizontal axis in Figure 8.1 denote the direction of motion of k). The following substantive results emerge:

1) *Existence of a poverty trap*

In the case of multiple equilibria in (8), there is a ranking of equilibria (see Figure 8.1) in terms of all per-capita variables, so that the locally stable 'low' equilibrium k_{low} is a poverty trap (or a Malthusian equilibrium), whereas the rich equilibrium k_{high} can be interpreted as a 'modern', post-transitional, regime. Barro and Sala-i-Martin (1995) argue that poverty traps typically arise as the result of the coexistence of regions of decreasing returns with regions of increasing returns. In (8) the poverty trap arises from the shape of the rate of population growth during the transition.

There is a second interesting case, namely when the m_1 curve is so low that the existence of k_{mid}, k_{high} is prevented. In this case the unique equilibrium k_1 is a 'very poor' equilibrium as, ceteris paribus, $k_1 < k_{low}$, and is stable. Therefore k_1 is still a poverty trap. This case can be considered historically prior to the three equilibria case, representing that time-window of the '*ancien régime*' stagnation during which accumulation was so low that the existence itself of a richer regime was impossible. Finally, there is a third case (not showed in Figure 8.1) when m_1 is so high that the existence of the poor equilibria k_{low}, k_{mid} is prevented, k_1 becomes a 'virtuous' modern equilibrium. This leads to the following remark concerning the non-ineluctability of modernisation:

Remark 2: The existence of a poverty trap suggests that the traditional view, which views *modernisation* (the industrial revolution and the DT) as an ineluctable process, is incorrect.

Once a poverty trap exists, a major problem is of course how to break out of it. A common feature of models such as (8) is that the 'escape' may only occur as a consequence of policies and/or external shocks, such as i) domestic policies/shocks leading to an increase in the saving rate (as documented in Strulik, 1997, 1999b), ii) increases in the technological parameters tuning both the height and/or the shape of the m_1 curve. Though such external events are not the only explanations of the 'escape', they certainly took place in history, and played a role in giving the 'big push' to investment in capital that allowed the escape from the Malthusian stagnation towards the Modern Regime. This is clearly expressed by Becker et al. (1990, S33): 'We believe that the West's primacy, which began in the XVII century was partly due to a "lucky" timing of technological and political changes in West'.

2) *Realistic convergence patterns*

Let us consider the rate of change of k in the region (k_{mid}, k_{large}) where escape from the trap has occurred. As Figure 8.2 shows, the rate of change of k, γ_k, increases from zero (at k_{mid}) up to a maximum at k^*, and then monotonically declines to zero again, as the system is attracted in the k_{high} equilibrium. The implications are noteworthy. In Solow's (1956) model γ_k monotonically declines to zero as k increases toward its equilibrium value.

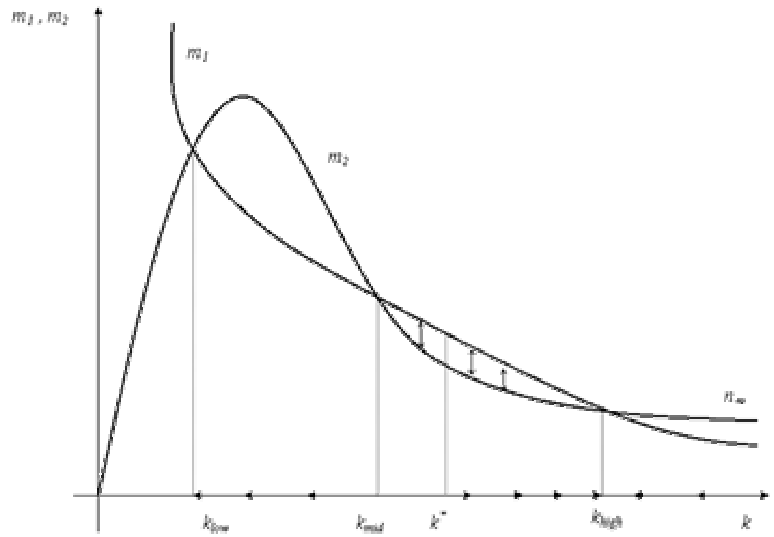


Figure 8.2 – Realistic (absolute) divergent/convergent patterns in the CRS Solow model with demographic transition

This has led to the controversial concept of absolute convergence (Barro and Sala-i-Martin, 1995, ch. 1): economies with lower per-capita capital are predicted to grow faster in per-capita terms, a fact which has often been denied on empirical grounds (quotation from Chu in Section 8.1). In model (8) economies which in the end escaped the Malthusian trap do not exhibit convergent paths. Consider two economies A, B which escaped the trap, i.e. that after some external shock entered the (k_{mid}, k_{high}) region with respective endowments $k_{poor} < k_{rich}$. (not shown on the figure). There is a whole region in the set (k_{mid}, k_{high}) in which the two economies initially diverge, i.e. the richer economy grows faster than the poorer one (the amplitude of such a region depends on the actual position of $k_{poor} < k_{rich}$). In other words richer countries become, in an initial phase (the temporal length of which may be quite large) even richer. Only at a later stage will the two economies enter a phase of convergent dynamics as in Solow's model. We argue that the present mechanism offers the simplest explanation for currently observed paths of rich versus developing countries. Such an explanation has the merit of being based on a huge body of demographic evidence, and therefore appears more convincing of alternative explanations, such as that offered in Barro and Sala-i-Martin (1995, ch. 1), which is based on the rather controversial coexistence of windows of increasing and decreasing returns. Therefore the following remark holds:

Remark 3: the diverging growth paths among countries observed over the past 50 years can be wholly compatible with Solow's model with DT, in that the length of the first phase after the escape from poverty can be quite large.

8.4.1. Adding Technological Progress

Let us consider exogenous technical progress at the constant rate α, combined with CRS Cobb–Douglas technology, i.e. $Y = Qe^{\alpha t} L(t)^{\beta} K(t)^{1-\beta}$. This leads to the following model in the per-capita variables $k = K/L$, $y = Y/L$

$$\begin{cases} \dfrac{\dot{k}}{k} = s\dfrac{y}{k} - \delta - n(y) \\ \dfrac{\dot{y}}{y} = \alpha + (1-\beta)\dfrac{\dot{k}}{k} \end{cases} \tag{9}$$

The previous system does not have non-trivial equilibria in the per-capita variables (if $\dot{k}/k = 0$, then $\dot{y}/y = \alpha > 0$). It has only a stable state of balanced growth. This fact would lead to unrealistic predictions in terms of very long term growth. In particular, because of the well documented long-term growth in technology, it would imply a sustained economic growth much before than when it has been observed. This result suggests that steady technological growth plus CRS cannot have been the rule in historical epochs. Hence, the hypothesis of CRS as the rule during very long-term economic development appears as very implausible. Moreover, the long-term rate of growth of per-capita variables in the balanced growth regime is the rate of labour augmenting technical progress $q = \alpha/\beta$, i.e., it is independent of the (long-term) rate of change of the population. In other words, the CRS model would lead to the conclusion that the demographic transition played no role in long-term growth. This suggests the following:

Remark 4: The CRS model is genetically unable to study the features of the DT and to explain the historical pattern of the economic 'take-off'.

8.5. DEMOGRAPHIC TRANSITION AND THE SOLOW MODEL UNDER DECREASING RETURNS TO SCALE

Motivated by the last remark of the previous section we consider now a model with Decreasing Returns to Scale (DRS). We acknowledge that the complexity of the DT would require much greater detail to take into account the changing economic environment. As suggested by both the historical evidence (Livi-Bacci, 1997, 1998; see also the discussion in Section 8.2.1) and the recent history of developing countries, a non-trivial investigation of the implications of the DT would imply the need to simultaneously endogenise all the parameters of Solow's model.[12]

Consideration of further realism in descriptive neoclassical models of the DT has led to several research efforts in very recent times. Strulik (1999a) considered a DRS Solow-type model for the demographic transition, saving transition (as suggested by both theoretical and empirical evidence) and fixed wage regulation (typical of developing countries). Unfortunately, the fixed wage regulation assumption, in which the workers are exposed to a forever constant wage, requires further assumptions to explain why they should experience a transition in their demographic patterns as per-capita income increases without increasing wages, and moreover the possibility of the Modern regime equilibrium is vitiated by a violation of the supply of labour constraint. Prskawetz et al. (2000) investigated, within a similar framework, the role of human capital as an additional factor in the production function in the presence of a 'transitional' dynamics of the rate of technological progress. However, their descriptive introduction of human capital in the production function is rather ad hoc. Strulik (1999b) considered the transition in a broader framework, with a microfounded approach to human capital, and a more general production function exhibiting both increasing and decreasing returns, while Strulik (2000) investigated the transition in a two-sector Solow-type economy.

The model that is presented here may be considered the neo-classical backbone of all the aforementioned contributions in that it is simpler while at the same time preserving the main dynamic features of richer formulations, such as Strulik (2000) and Praskawetz et al. (2000). Moreover it does not suffer the contradiction of Strulik's (1999a).

Our model assumes the following DRS production function:

$$Y = Qe^{\alpha t} L(t)^{\beta} K(t)^{\gamma} \qquad Q > 0 \;\; ; \;\; \beta + \gamma - 1 < 0; \alpha > 0 \tag{10}$$

We moreover assume 'transitional' dynamics (over the same time window) of: *a*) the rate of change of the population according to a positive humped function of per-capita income, *b*) the saving rate, which is assumed to be an increasing logistic function of y (as in Strulik, 1999a, 1999b, 2000; Praskawetz et al., 2000); *c*) the rate of change of technical progress α, taken as an increasing logistic function of y, as in Prskawetz et al. (2000). Standard Solow-type assumptions and straightforward manipulations lead to the following formulation in the per-capita variables $x = K/P$, $y = Y/P$ where P is the total population size (or the total supply of labour)[13]

$$\frac{\dot{x}}{x} = s(y)\frac{y}{x} - \delta - n(y) \tag{11a}$$

$$\frac{\dot{y}}{y} = \alpha(y) + \gamma\frac{\dot{x}}{x} - (1-\beta-\gamma)n(y) \tag{11b}$$

It is convenient for analysis to consider the variables $U = y/x$, and y, obtaining:

$$\frac{\dot{U}}{U} = \alpha(y) + (1-\gamma)\delta + \beta n(y) - (1-\gamma)s(y)U \tag{12a}$$

$$\frac{\dot{y}}{y} = \alpha(y) - \gamma\delta + \gamma s(y)U - (1-\beta)n(y) \tag{12b}$$

In order to gain a broad understanding of the results of the present section, it is worth making the following remark concerning the basic neo-classical DRS model (i.e. model (12) for exogenously determined s, n, α):

Remark 5 (neo-classical model with DRS**)**. Under constant s,n,α model (12) has a unique globally asymptotically stable state of balanced evolution in per-capita variables. The long term rate of change of per-capita variables is:

$$q = \frac{\alpha - (1-\beta-\gamma)n}{(1-\gamma)}$$

In particular balanced growth occurs for $\alpha - (1-\beta-\gamma)n > 0$, whereas balanced decay occurs in the opposite case. In the absence of technical progress ($\alpha = 0$) the long term outcome is always balanced decay at the rate $q = -\left((1-\beta-\gamma)n\right)/(1-\gamma)$.

Let us now investigate (12) in a hierarchical manner, by starting from the basic situation in which only population is endogenous, whereas s, α are taken as constant. Subsequently, in addition to the endogenous population, also the cases of, respectively, endogenous saving (Section 8.5.2) and endogenous technical progress changes will be investigated (Section 8.5.3).

8.5.1. Only Population Growth is Endogenous

Equilibrium analysis gives the non-trivial nullclines

$$U = \frac{\alpha + (1-\gamma)\delta + \beta n(y)}{(1-\gamma)s} = h_1(y) \quad ; \quad U = \frac{\gamma\delta - \alpha + (1-\beta)n(y)}{\gamma s} = h_2(y) \tag{13}$$

The curves $h_1(y)$, $h_2(y)$ inherit the humped shape of $n(y)$. Let us consider the standard case of stationary population at both the beginning and end of the transition: $n(0) = n(\infty) = 0$.[14] In this case $h_1(0) > h_2(0)$, and $h_1(\infty) > h_2(\infty)$. There are two possibilities: *a*) the curve $h_1(y)$ always lies above $h_2(y)$ (no equilibria), or *b*) the curve $h_2(y)$ intersects twice $h_1(y)$.

Case a) occurs for $\alpha > (1-\beta-\gamma)n_{Max}$, where n_{Max} is the maximal growth rate attained by the population during the DT, whereas b) occurs in the opposite case, $\alpha > (1-\beta-\gamma)n_{Max}$. Therefore no equilibria exist in case *a*), whereas exactly two equilibria exist in case *b*). In this latter case let us denote the equilibria by E_1,E_2, where E_1 is the 'poor' equilibrium with smaller per-capita income. It is easy to show (see Appendix) that E_1 is always locally asymptotically stable (LAS), whereas E_2 is always unstable. The nullclines and the directions of motion of system (12) are drawn in Figure 8.3 for case *b*). To fully understand the dynamics of the system it is also useful to look for states of balanced growth, i.e. asymptotic states characterised by exponential growth of per-capita variables at the same constant rate q, and therefore by an asymptotic constant ratio $U^* = H$. This leads to the following system in the quantities (H,q)

$$\alpha + (1-\gamma)\delta + \beta n(\infty) - (1-\gamma)sH = 0 \tag{14a}$$

$$q = \alpha - \gamma\delta + \gamma sH - (1-\beta)n(\infty) \tag{14b}$$

giving:

$$q = \frac{\alpha - (1-\beta-\gamma)n(\infty)}{1-\gamma} \quad ; \quad H = \frac{\alpha + (1-\gamma)\delta + \beta n(\infty)}{(1-\gamma)s} \tag{15}$$

Examination of Figure 8.3 shows that the state of balanced growth (15) is LAS. Note that the asymptotically constant portion of the h_1 curve represents the radius of the balanced growth state.[15]

Let us summarise the main results by the following:

Proposition 2. In the neo-classical model (12) with endogenous population dynamics, given the technological ratio γ/β, two main outcomes are possible: *a*) the rate of change of technology (α) is so large[16] as to always absorb the population bulge observed during the transition. In this case the system always attains a long-term state of balanced growth, at the rate $q = \left(\alpha - (1-\beta-\gamma)n(\infty)\right)/(1-\gamma)$; *b*) the growth of technology is not sufficiently fast to absorb the population bulge. In this case two equilibria (E_1 and E_2) are possible and, depending on initial conditions, the economy will be attracted in the Malthusian poverty trap E_1, or will attain a long-term state of balanced growth.

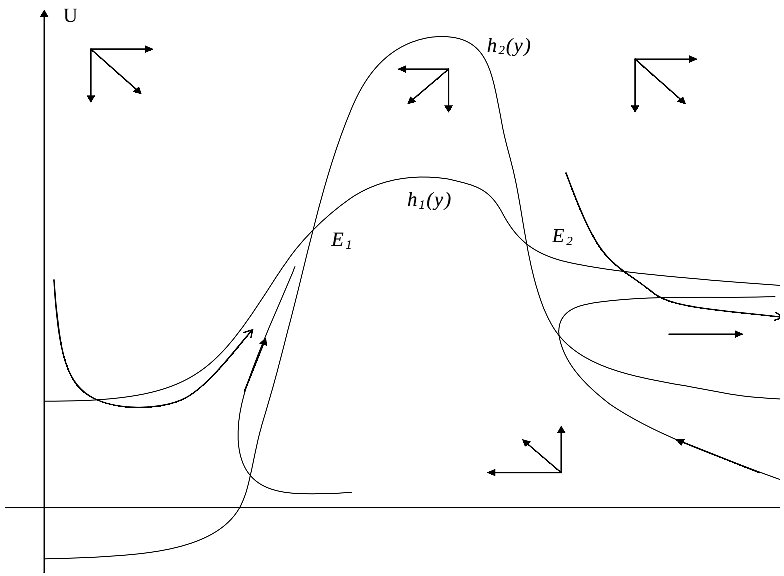


Figure 8.3 – Equilibria and direction of motion for the neo-classical DRS model with DT in case b) (two equilibria)

Remark 6. The previous result suggests that, contrary to the CRS case, the DRS framework allows: a) the simultaneous coexistence of stable ‘historical’ steady states with ‘modern’ long-term stable states of balanced growth, characterised by different basins of attraction. That is: the DRS model offers a unified explanation of both Malthusian stagnation and modern growth. b) that only in the DRS framework the population dynamics affects per-capita income growth. This shows to what extent the DT may have been crucial in allowing modern regimes with fast economic growth.

8.5.2. Endogenous Population and Saving

We now move on to study the joint dynamic effects of population transition and ‘saving transition’.[17]

The non-trivial nullclines are:

$$U=\frac{\alpha+(1-\gamma)\delta+\beta n(y)}{(1-\gamma)s(y)}=h_3(y)\ ;\ \ U=\frac{\gamma\delta-\alpha+(1-\beta)n(y)}{\gamma s(y)}=h_4(y) \quad (16)$$

Let us summarise the main results by the following:

Proposition 2. In the neo-classical model (12) with endogenous population dynamics, given the technological ratio γ/β, two main outcomes are possible: *a*) the rate of change of technology (α) is so large[18] as to always absorb the population bulge observed during the transition. In this case the system always attains a long-term state of balanced growth, at the rate $q=(\alpha-(1-\beta-\gamma)n(\infty))/(1-\gamma)$; *b*) the growth of technology is not sufficiently fast to absorb the population bulge. In this case two equilibria (E_1 and E_2) are possible and, depending on initial conditions, the economy will be attracted in the Malthusian poverty trap E_1, or will attain a long-term state of balanced growth.

which are not necessarily humped, as in the previous case. Under the standard assumption $n(0) = n(\infty) = 0$, it holds that $h_3(0) > h_4(0)$, and $h_3(\infty) > h_4(\infty)$, with $h_3(0) > h_3(\infty)$, and $h_4(0) > h_4(\infty)$.There are two interesting cases depending on whether the initial population growth has been faster, or slower, compared tothe saving rate growth. In the first case the curves $h_3(y)$, $h_4(y)$ are, in most situations, one-humped, as in Section 8.5.1, whereas in the second case both curves will be, in most situations, monotonically decreasing in *y*. In terms of equilibria this implies that again, as in the previous section, two main cases are possible: *a*) no equilibria, when the rate of exogenous technical progress is much larger compared to the maximal rate of increase of the population during the transition, and *b*) two equilibria in the opposite case. The dynamic analysis shows results quite similar to those of Section 8.5.1 (the poor Malthusian state is usually LAS, whereas the second equilibrium E_2 is unstable) with a major novelty, e.g. the appearance of stable oscillations around the Malthusian equilibrium. Such stable oscillations appear through a Hopf bifurcation of the Malthusian equilibrium (see Appendix). The major substantive results are summarised in the following:

Proposition 3. In the neo-classical DRS model with endogenous population and ‘saving transition’, in addition to the outcome formulated in Proposition 2, also the following outcome is possible: the stagnant Malthusian dynamics around the Malthusian poverty trap E_1 may occur through stationarity, or through stable oscillations. The appearance of stable oscillations is the consequence of a quick change in the patterns of saving at the beginning of the transition, which locally (but only locally) destabilises the Malthusian equilibrium.

The previous proposition suggests that those persistent oscillations that scholars in economic history have usually explained through purely exogenous arguments, plague or famine crises for instance, could have been just part of the story. In fact the economic system was potentially able to lead to purely endogenous oscillations around the Malthusian stagnation equilibrium, for instance via the mechanism embedded in the model of this section.

In sum we have shown that the DRS framework is capable of capturing two major features of long-term demo-economic evolution: the existence of a Malthusian, possibly oscillatory, regime and the crucial role played by the DT in allowing modern regimes with fast economic growth.

8.5.3. The Joint Role of Endogenous Productivity Changes and Population

Major studies in economic history (Bairoch, 1973) have amply documented that technical progress has not been constant in the long term and its time profile can be represented by an s-shaped curve, mirroring the initial phase of very slow increase followed by substantial acceleration from the beginning of the industrial revolution, and by a decrease in the pace of growth in the more recent phase.[19] The analysis may lead to two distinct cases depending on whether the population increase temporally followed, rather than preceded, the increase in technology. For western countries the pace of technological growth is known to have started to increase long before the onset of the DT. In this case we again have the coexistence of a locally stable Malthusian equilibrium, with an unstable intermediate equilibrium, with a locally stable regime of balanced growth. Compared to Section 8.5.2 no persistent historical oscillations are possible (see Appendix). In the opposite case, in which the DT anticipated the technological take-off, and which is relevant for many developing economies, a third equilibrium may appear (oscillations are again ruled out) making the mechanisms of escape from the poverty trap richer than the previous cases. This result was also obtained by Prskawetz et al. (2000) but in a different model.

To sum up, our models presented above have ‘benchmarks’ with which to compare all the models generating multiple equilibria (e.g. the models of Becker et al., 1990 and Lucas, 1998). Although the two latter models are completely different from ours in that they are *i*) microfounded according to a different history with respect to that described in Section 8.3.1, and *ii*) are centred on the role of human capital accumulation, they share with our model the focus on the population dynamics in a Malthusian and a Modern era, and in particular the same interpretation as follows: the Malthusian and Modern eras are different steady states of the same model.

8.6. CONCLUDING REMARKS

The present paper, by re-emphasizing the endogeneity of the supply of labour as in the classical economics, shows that also the standard neoclassical Solowian model, can offer insights into fundamental aspects of modern growth. We critically discuss several aspects related to the modelling of the Demographic Transition, offering a microfounded explanation of the latter. In particular the basic Solow 1956 Constant Returns to Scale (CRS) model, having been enriched with an assumption on the dynamics of the supply of labour mimicking the DT, offers interesting insights into the interrelationships between population dynamics and poverty traps, and a simple, but interesting, generalisation of the widely debated notion of convergence. Moreover when also (exogenous) technical progress is considered we argue that 1) the CRS model is genetically unable to study the features of the DT and to explain the historical pattern of the economic ‘take-off’; 2) by contrast, our neo-classical Decreasing Returns to Scale framework is capable of showing the coexistence of a stable poor ‘Malthusian stagnation’ (possibly oscillatory) equilibrium with stable states of ‘modern’ long-term balanced growth, as historically observed, and especially of showing to what extent the DT may have been crucial in allowing modern regimes with fast economic growth. This feature suggests that DRS models, rather than CRS, are optimal candidates as unified models of ‘very-long’ term economic growth, and that the enrichment of the neoclassical growth models with themes belonging to classical economics, namely the supply of labour dynamics and decreasing returns, can provide an interpretative tool of the DT and economic ‘take-off’.

APPENDIX: LOCAL STABILITY ANALYSIS OF EQUILIBRIA IN THE NEO-CLASSICAL MODEL WITH DEMOGRAPHIC TRANSITION AND DECREASING RETURNS TO SCALE

The Jacobian of (12) evaluated at equilibrium

$$J=\begin{pmatrix} -(1-\gamma)s(y)U & U\left(\alpha'(y)+\beta n'(y)-(1-\gamma)s'(y)U\right) \\ \gamma y s(y) & y\left(\alpha'(y)-(1-\beta)n'(y)+\gamma U s'(y)\right) \end{pmatrix}$$

A) In the model with only endogenous population (Section 8.5.1) we obtain

$$Tr(J) = -(1-\gamma)sU_1 - (1-\beta)n'(y)y\,,\ Det(J) = sUn'(y)y(1-\beta-\gamma) > 0\,.$$

Since $n'(y) > 0$ in the Malthusian equilibrium E_1, then $Tr(J_1) < 0$, whereas $Det(J_1) = sUn'(y_1)y_1(1-\beta-\gamma) > 0$, showing that E_1 is always LAS. Similarly, since $n'(y) < 0$ at E_2, E_2 is always unstable.

B) In the model with endogenous population and saving (Section 8.5.2) we obtain:

$$Tr(J) = -(1-\gamma)s(y)U + y\left(-(1-\beta)n'(y) + \gamma Us'(y)\right).$$

Moreover: $Det(J) = ys(y)U(1-\beta-\gamma)n'(y) > 0$. The last expression shows that a Hopf bifurcation occurs for some parameter constellations at the Malthusian equilibrium. In fact the Trace may become negative for some parameter constellations, especially when the saving rate increases rapidly at the beginning of the transition era, whereas $Det(J_1) > 0$.

C) In the presence of endogenous population and technological change (Section 8.5.3) at the Malthusian equilibrium E_1 the following holds:
Let us summarise the main results by the following:

Proposition 2. In the neo-classical model (12) with endogenous population dynamics, given the technological ratio γ/β, two main outcomes are possible: *a*) the rate of change of technology (α) is so large[20] as to always absorb the population bulge observed during the transition. In this case the system always attains a long-term state of balanced growth, at the rate $q = (\alpha - (1-\beta-\gamma)n(\infty))/(1-\gamma)$; *b*) the growth of technology is not sufficiently fast to absorb the population bulge. In this case two equilibria (E_1 and E_2) are possible and, depending on initial conditions, the economy will be attracted in the Malthusian poverty trap E_1, or will attain a long-term state of balanced growth.

$$Tr(J) = -(1-\gamma)sU_1 - y\left((1-\beta)n'(y_1) - \alpha'(y_1)\right),$$

and

$$Det(J) = sU_1 y_1\left((1-\beta-\gamma)n'(y_1) - \alpha'(y_1)\right).$$

This implies that no Hopf bifurcation is possible at the Malthusian equilibrium.

NOTES

1. In this section we do not consider other features not directly related to fertility and mortality, such as technological progress dynamics and the saving 'transition', that the recent literature – reviewed in the next section – embodies in the more general historical process connecting demographic transition, industrial revolution and the modern growth era. Some possible roles played by such phenomena are taken into account in the third part of this work.
2. In some cases however there is also evidence (Dyson and Murphy, 1985) that fertility temporarily rose, thanks to increased well-being, before starting to decline.
3. The same terminology may express a different concept in other authors: e,g Kögel and Prskawetz (2001, p. 2) claim that: '... economy will be trapped in a situation where no sustained growth of per capita income can occur. This trap is commonly labeled Malthusian trap' and 'we label the simultaneous take-off in economic growth and population growth as escape from the Malthusian trap'. In our opinion these definitions are restrictive: they reflect the emphasis on the 'forever sustained' exponential growth which pervades all the endogenous growth theory.
4. Finally, note that from the demoeconomic perspective, the Modern Regime is sometimes defined also as Post-Classical, Anti-Malthusian, Beckerian.
5. Some issues associated with this recent literature are reviewed in Fanti and Manfredi (2001).
6. Notice that this utility function, by assuming $\gamma = \eta < 1$, shows a high elasticity of substitution between consumption and children (e): indeed $e = 1/\eta > 1$ and the elasticity is decreasing with η (at the limit when η tends to one, the elasticity tends to become unitary as in the Cobb–Douglas case).
7. This argument is developed by Strulik (1997, 1999a) who models the population birth (*b*) and death (μ) rates during the transitional regime by means of two asynchronously decreasing logistic-alike functions of per-capita income *y*. Moreover, an attempt to provide a theoretical justification based on an overlapping generation argument is also given in Strulik (1999a).
8. More general explanations seek to embody both types of effects (Retherford and Palmore, 1983).
9. Rosero-Bixby and Casterline (1993) suggests three main types of effects which may lead people to switch to a different group: not only information flows, but also demonstration effects, based on the experience of other people which evidence the benefit of the transition, and changes in normative contexts.
10. Equations (6a)–(6b) can also be solved analytically.
11. We will not be concerned with post-transitional (or 'second' demographic transition) phenomena (of which Italy and Spain in the past 15 years are major examples), by which the rate of population change could even become sharply negative.
12. An example is the rate of growth of technology, which has not been constant since historical times, but has rather experienced a historical evolution from an initial value close to zero, followed by a very slow increase, before the take-off of the modern era, which is complexly correlated with growing per-capita income and knowledge. Similarly, the constancy of the parameters tuning the relative role of capital and labour in the production function is, at best, a simplistic approximation, and so on.
13. The distinction between variables $y = Y/P$ and the traditional Y/L is formally unnecessary here but it is useful in more general contexts.
14. Some differences would arise if $n(0) > 0$ and/or $n(\infty) > 0$ (or even $n(\infty) < 0$, according to the second demographic transition). As the purpose of this paper is not taxonomical we do not consider them here.

15. As regards the basins of attraction of the balanced growth state, note that the region to the right of the unstable equilibrium E_2 and above the curve h_1 is positively invariant, as is the region to the right of E_2 between the two curves.
16. Alternatively one could reason in terms of the composite technological parameter $\alpha/(1-\beta-\gamma)$ versus the rate of population growth.
17. The latter has been documented empirically for some Western countries by Strulik (1997), who has also given theoretical support to the fact that the mortality decline observed during the DT could lead to an increase in the saving rate as well.
18. Alternatively one could reason in terms of the composite technological parameter $\alpha/(1-\beta-\gamma)$ versus the rate of population growth.
19. For instance, as stated above, Prskawetz et al. (2000) used a logistic function.
20. Alternatively one could reason in terms of the composite technological parameter $\alpha/(1-\beta-\gamma)$ versus the rate of population growth.

9. Human capital formation in the new growth theory: the role of 'social factors'

Maria Rosaria Carillo

9.1. INTRODUCTION

First-generation endogenous growth models, assuming human capital accumulation as a major engine of growth, have grounded their analysis on the Beckerian model of human capital, where homogeneous agents in the presence of perfectly competitive markets forgo leisure and current income in order to increase their knowledge and obtain a higher future income. This approach envisages no role in the creation of human capital for any of the phenomena tied to an individual's social behaviour such as 'peer effect', 'direct knowledge transmission', 'status-ranking' of occupations, 'network relationship', and so on, although the importance of such social phenomena for individual formation has been widely recognized by the literature on human capital.[1] Probably, behind the recognition that human capital has the distinctive feature of producing a large amount of externalities (Lucas, 1988) lies the idea that direct social relations among agents themselves create knowledge. Nevertheless, this phenomenon has not been explicitly investigated, and the mechanisms by which the externalities are generated remain entirely unexplained.

In recent years, a class of endogenous growth models have analysed in depth how knowledge is formed and transmitted among individuals to give rise to externalities. In doing so, they have highlighted the role played by social relations in the creation of human capital, by assuming that the latter is formed not only through an educational activity, but also through the relations that arise among individuals. More specifically, they show that social factors are further channels for the transmission of knowledge which also modify its use and desirability.

The growth role of social factors has been largely analysed by the New Growth Theory, without limiting the analysis to the effects on human capital formation. Cole et all (1992) and Corneo and Jeanne (1997), for example, have analysed the effects of status-seeking behaviour on wealth accumulation and on the saving rate. Temple and Johnson (1998), using the Adelman and Morris data base, carry out an empirical analysis to test whether 'social capital' matters in determining economic performance, thereby confirming this hypothesis. Also Knack and Keefer (1997) found evidence that trust and cooperation are associated with a higher economic growth rate while Zak and Knack (2001), assuming that trust reduces transaction costs, show that high trust societies have a higher investment rate and produce more output than low trust societies.

A large number of papers follow this strand of literature and there is ever growing attention towards this field of research. In this paper I shall consider only a particular aspect of this wide theme: the growth effects of social factors via their influence on the accumulation of human capital. Before analysing this theme, I shall try to define what are social factors and the nature of the relation existing between them and human capital. An initial problem is the lack of clarity over the definition of social factors. Moreover, it will be apparent that also the nature of the relation between human capital and the latter factors is not clear. Several authors hypothesise that they directly affect the human capital accumulation process, which occurs with important feedback effects (Coleman, 1988). However, others assume that social factors influence human capital accumulation only indirectly, since they are productive factors in the aggregate production function which are complementary to, or substitutes of human capital (Glaeser et al., 2000; Iyigun et al., 2001).

Analysis of the relation between human capital and social factors is an important theoretical aspect, since it leads to very interesting results such as the persistence of heterogeneity of individuals and the possibility that social classes[2] and a wage structure reflecting not only differences in productivity, but also the social organisation, may emerge. Both of them are themes largely analysed by Classical Theory.

Another important feature of this class of models which is a resurgence of a classical theme, is the assumption that the behaviour of a rational agent also depends on some extra economic factors almost always related to social relations with other agents. For example, in Fershtman et al. (1996) agents take care of their social position, and the level of human capital is chosen by considering also the effect on their social reputation, in Gradstein and Justman (2000) agents undertake an education activity for conformism, while in Galor and Tsiddon (1994) parents transfer their human capital to children for altruism.

The social aspects underlying the agents' economic behaviour are well known by classical authors. Smith (1776) in his analysis of the nature of wage highlighted the fact that the reward structure of different occupations

also reflects the reputation associated to each. This happens since individuals take care of their social relations, and the total reward of an occupation is affected also by the relative position in the social ranking obtainable with that occupation. 'First the wages of labour vary with…the honourableness or dishonourableness of the employment. … Honour makes a great part of the reward of all honourable professions' (Smith 1776, p. 202). Again he wrote 'The public admiration … makes a considerable part of total reward in the professions of physic, a still greater perhaps in that of law; in poetry and philosophy it makes almost the whole' (Ibidem, p. 209).

This hypothesis about individual behaviour has been adopted by several strands of literature, yet the inclusion in an endogenous growth framework opens up further lines of research, since it makes it possible to analyse how social interdependence can interact with the growth process: according to which behaviour will prevail and which relations between agents will become stronger or weaker, it is possible to predict the evolution of 'types' of individuals within the population. Therefore this approach can lead naturally to an evolutionary analysis of economies and of their economic performance.[3] From these considerations, it is apparent that the possible developments of this approach are considerable.

This chapter, which will review this strand of literature, is organized as follows. The second section discusses the concepts of *social interactions*, *social capital*, *culture* and *ideology*, and *social status* concern, all of which indicate the effects of social factors on human capital formation and its diffusion among individuals. The third section surveys the analytical methods proposed to include in the economic analysis the effects of social factors on economic growth via their influence on the labour factor. Here it will be argued that a general framework, which can encompass in the economic analysis the effects of social factors, is still lacking. The fourth section contains some proposals for a solution to this problem. The chapter concludes with some brief remarks.

9.2. SOCIAL FACTORS AND HUMAN CAPITAL IN NEW GROWTH THEORY

Although there is no general agreement on the nature of the relation between human capital and social factors, the most generally accepted idea is that social factors affect human capital accumulation through different channels. The most important is *direct social relations* among agents, since individual human capital can be acquired not only through an educational activity undertaken in school, but also from other individuals with whom they have social contacts. Normally, individuals from whom knowledge can be acquired are agents whose services can be bought in the market, but this is not always true. They can be relatives and parents, for example, who transmit their knowledge without receiving recompense for it. In such cases, knowledge is transmitted because of the relations among individuals in the absence of a market and without a price being formed for it.

Another channel is the *culture*, the *norms* and *beliefs* that characterise a community. The latter constitute a considerable part of the human capital that individuals possess, and are transmitted[4] to all the members of the community only because they belong to it.

Finally, social factors may influence the level of education since they modify the incentives to acquire new knowledge. An example is the case when individuals desire more human capital not to earn a higher income, but to acquire a higher social status or conform to their group. The incentives in this case lie in the social relations among individuals who attend to their relative social position.

Second-generation growth models have included these factors in their analysis of human capital accumulation. However, the set of analytical categories used until now, denoted here as social factors, is rather heterogeneous. There are in fact models which have focused mainly on the effects of *social interactions* (Benabou, 1996 and Durlauf, 1996), where the term refers principally to interactions among single individuals or between these and reference groups. Others have instead emphasised the role of *social capital*, by which is meant a broad and heterogeneous set of phenomena including the social norms and institutions that characterize a society (Coleman, 1988). Finally, some have included *culture* and *ideology* (Cozzi, 1998; North, 1981; Iyigun et al., 2001) among society's 'social assets' which influence the formation of human capital. There is, therefore, a plethora of concepts and analytical categories which seem difficult to sum up in a single term. However, all of them relate to phenomena that spring from direct interdependence among individuals, 'direct' in the sense that it is not mediated by market mechanisms. In what follows, I shall analyse the various analytical categories used to grasp the effects exerted by social factors on agent formation and on economic performance.

9.2.1. Social Interactions

Social interactions constitute one of the most widely used analytical categories to describe the effects of social factors on the labour supply in terms of efficiency units. Their effects on the growth process have been analysed mainly by Benabou (1996) and Durlauf (1996).

A clear definition of social interactions has been provided by Brock and Durlauf (2001), as follows: 'By social interactions we refer to the idea that

the utility or pay-off an individual receives from a given action depends directly on the choices of others in that individual's reference group…' (Ibidem, p. 235). These are therefore relations among individuals of economic importance, because the action of one agent influences the chosen action of another agent with whom s/he is directly or indirectly linked. The main hypothesis is that agents influence each other through their actions and not through other media like, for example, directly exchanged information.[5] Another crucial hypothesis is that this influence comes about directly, in the sense that it does not operate through the market: the individual modifies her/his rational choice simply by observing the actions of other agents.

Generally the literature distinguishes between local and global interactions. Global interactions arise when an agent is able to interact with any other agent in the economy. Local interactions are cases in which agents interact only with some specific group of agents. In the latter case, a set of neighbours must be defined, and the notion of social distance between agents is also required.

One major difficulty concerns the way in which the effects of such interactions can be included in the choice problem. Three different approaches exist in the literature: (*i*) it is assumed that social interactions modify the constraints under which the rational choice is made; (*ii*) they influence the formation of expectations; (*iii*) or modify the preference structure.

Constraint interactions occur when an agent's action modifies the choice set of other agents. An example is provided by the congestion or spillover effects due to knowledge diffusion. Interactions through expectation formation are well known, and their analysis pervades information theory. Economic analysis assumes that an agent forming expectations may seek to draw lessons from observation of the actions chosen by others. Preference interactions occur when an agent's preference over the alternative in a choice set depends on the actions chosen by other agents, examples being provided by conformism, jealousy and envy. This way of modelling the interaction among agents can be traced back to Smith (1759) who recognised that agents interact through preferences[6] and that the intensity of interactions varies according to the strength of the relationship.[7] Also this last idea is embedded in social interaction models, since another key assumption is that the intensity of the effects of social interactions depends on the strength of the relationship among agents, usually captured by the social distance function, used to gain a measure of the social 'nearness' between two or more agents and the degree of interaction between them.[8]

9.2.2. Social Capital

Social capital is another of the concepts used to specify the influence exerted by social factors on the behaviour of economic agents. Providing a definition of the term is difficult, since in this case there is no general agreement on its meaning in the literature. Moreover, the concept often overlaps with that of social interactions. One of the best known definitions is provided by Putnam (1993, p. 167): ‘social capital ... refers to features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions’.

A similar definition is provided by Coleman, to the effect that: ‘Social capital is defined by its functions. … (it) consists of some aspects of social structures, that facilitates certain actions of actors. Like other forms of capital, social capital is productive making possible the achievement of certain ends that in its absence would not be possible’ (1988, p. S98).

According to this second definition, social capital consists of the mechanisms that facilitate the coordination of individual actions so that a superior outcome is achieved. As Durlauf (1999a) and Woolcock (2000) have noted, the problem with this type of definition is that it concentrates on the possible effects while ignoring the mechanisms that create social capital. It thus confuses a positive outcome with what has made that outcome possible.

A definition which overcomes this problem has been suggested by Durlauf (1999a), for whom social capital is ‘the influence which the characteristics and behaviours of one’s reference group have on one’s assessments of alternative courses of behaviour’ (Ibidem, p. 2). Rather than emphasising the ‘productivity’ of social capital, this definition stresses the ‘sociality’ of individual behaviour. It highlights the important role of non-market relationships in determining individual and collective behaviour, allowing the sources of social capital to be separated analytically from its consequences.

Besides the difficulty of coming up with a general definition of social capital, there is also the problem of defining the forms that it assumes. Woolcock (2000) proposes a scheme in which social capital may assume four dimensions: (*i*) the extent of horizontal associations; (*ii*) the nature of social ties within communities (the degree of trust, peer effect, etc.); (*iii*) the nature of the relation between civil society and the state and (*iv*) the quality of the governing institutions.

Coleman (1988) proposes a similar scheme, although he places greater stress on the role of collective norms, concluding that social capital assumes three forms: (*i*) obligations and trustworthiness of structures; (*ii*) information channels; (*iii*) norms and effective sanctions. Yet these classifications, too, suffer from the shortcoming of confusing social capital with its possible positive effects. Moreover, many of the concepts outlined above (social norms, for instance, or the quality of the governing institutions) have already been analysed in the literature without it being found necessary to introduce a new analytical category.

Although social capital may be a useful concept insofar as it underlines the sociality of individual action, and the effect of this on economic choices, it is either ill-defined or redundant. The definition proposed by Durlauf (1999a) deals with the former problem but makes the concept of social capital very similar to that of social interactions. Involved once again are relations among agents not mediated by the market; only that in the case of social capital the role of the reference group is more stressed.

9.2.3. Culture and Ideology

The concepts of culture and ideology (or religion) are two further analytical categories introduced by the new growth literature to analyse the effects of social factors on the creation of new knowledge and on economic performance (Casson, 1993; Cozzi, 1998; Gray, 1996; Iyigun et al., 2001; Lazear, 1999).

Culture is defined by anthropologists in a variety of ways, but whatever the definition, it usually includes some notion of shared values, beliefs, customs, rituals, language, and so on. Some authors (Gray, 1996) hypothesise that culture is a public good that affects the propensity of agents for various economic activities. Cozzi (1998) suggests that culture is a ‘social asset’ that increases the productivity of labour measured in terms of efficiency units, and which also accelerates the pace of technological innovation. The reason is that, although culture does not have an immediate productive use, it shapes individuals’ behaviour and thereby their productive capacity.

Another analytical category often used is ideology (or religion) (North, 1981, Iyigun et al., 2001). According to some authors (Sacerdote and Glaeser, 2001, North, 1981), ideology is a particular form of knowledge which enables generalizations to be made about the environment within which agents operate and completes the factual information that they possess. More specifically, by embracing an ideology, individuals increase their ability to acquire knowledge, and this affects positively the productivity of labour factor.

The positive relation between ideology and labour productivity has been hypothesised by several authors. Rosenberg and Birdzel (1986) point out that the development of a moral system commensurate with wealth and capital accumulation can be traced to the Calvinist Reformation of Protestantism.[9]

Franke et al. (1991) and Gray (1996) find evidence for a positive correlation between certain Confucian values and economic growth in samples comprising both Western and Asian countries. Iyigun et al. (2001) embed this idea in an endogenous growth model where ideology and education are substitutes and interact to influence technological progress.

Although the concept of ideology is useful since it unifies in a single term phenomena such as beliefs, moral and political attitudes that influence the behaviour of an individual, no significant difference between this and the concept of culture seems to emerge. Both refer to a particular type of knowledge which is shared by a multitude of persons, and have a pronounced normative content. Furthermore, they do not seem to differ greatly from the concept of social capital.

9.2.4. Social Status and Conformism

Social status is a ranking of individuals (or group of individuals) in a given society, based on their traits, occupation, consumption, assets and actions (Weiss and Fershtman, 1998). The sociological literature (Davis and Moore, 1945; Treiman, 1977) has shown that high social status is usually awarded to wealthier individuals and to those who have an occupation requiring a high level of human capital.

The economic implication of this social phenomenon has long been recognised by economists,[10] who have largely analysed also the implications for growth.[11] The influence of social status on growth has been assigned principally to its effects on the saving rate (Cole et al., 1992; Corneo and Jeanne, 1997) and on the demand for positional goods (Funk, 1996; Hirsh, 1976), while only few have recognised its growth effect via the influence on human capital (Fershtman et al., 1996). In the model of Fershtman at al. (1996) agents attend to their social status, obtainable by undertaking an occupation that gives high social prestige. Following Smith (1776), they assume that the latter is an attribute of those occupations which require high human capital. Therefore, the demand for social status may constitute a strong social incentive for the accumulation of human capital, which is added to the monetary incentive.

On the existence of this positive relation between human capital accumulation and concern for rank there is general agreement, but it is not so clear why individuals are concerned about their relative social position. According to Postlewaite (1998), the rank concern arises instrumentally because relative standing influences the consumption level. As a matter of fact, because of market imperfections not all goods or services can be acquired through the market. When the allocation of some goods or services such as information, education or other does not occur via the market, high

social status allows a high level of consumption of such goods to be achieved. This implies that the demand for social status derives from the presence of social interactions by which exchanges of non-marketable goods and services occur. Hence it is only another aspect of the more general phenomenon of non-market interactions.

However, the existence of decisions that affect consumption but are not mediated by price mechanisms can also lead to another type of behaviour, in contrast to the concern for status: the desire to conform. A conformist behaviour may emerge especially when there are activities that are undertaken in groups. It refers to an inclination of an individual to behave like the other agents belonging to his/her reference group. Also in relation to this different mode of social interaction the effects on human capital formation and on economic performance have been analysed. Gradstein and Justman (2000) propose a model where individuals gain utility from conformist behaviour by reducing the social distance between themselves and their reference group. They show that such behaviour may have perverse effects on growth because it may reduce the returns on investment in education.

At the end of this discussion, it seems clear that most of the concepts presented here suffer from a lack of definitional clarity. Moreover, the differences between them seem to be quite marginal. Social capital, for example, is a result of interactions among agents: the trust-degree, which is a dimension of it, is only a form assumed by social interactions. Also culture and ideology or rank concern are only particular results of the latter. All the concepts presented here are only different aspects of the same phenomenon: the social exchange of knowledge, information, etc. which occurs among agents, *social* because it is not mediated by the market and which for this reason the standard economic model, if not appropriately amended, is unable to capture.

9.3. ANALYTICAL METHODS TO ENCOMPASS SOCIAL FACTORS IN ENDOGENOUS GROWTH MODELS

Besides of the above-cited problems, analysis of the effects of social factors also suffers from a lack of a generally accepted analytical framework. It is not an easy task to encompass social factors in economic analysis given the fact that some of them refer to aggregate concepts. Indeed literature has almost universally viewed social capital, culture, norms and beliefs as community level attributes, but since economic models are based on decision maker agents, aggregate definitions may impede the inclusion of social factors in the economic framework. In spite of such difficulties, there have been several attempts that can be summarised in two analytical strategies: one where social factors modify the constraints under which the optimal choice is made, and another, more general strategy, where they modify the objective function or the preference structure.

9.3.1. Models where Social Factors Modify Constraints

An initial example of the adoption of this kind of strategy can be found in models where social factors modify the rewards structure by favouring or reducing human capital accumulation rate. Papers by Acemoglu (1995), Baumol (1990) and Murphy et al. (1991), for example, have analysed the effects of social factors on the allocation of talents, showing that if these make rent-seeking activities more remunerative than productive ones, the economy grows at a lower rate.[12] In all these models social factors influence the individual human capital indirectly, since they affect the relative convenience of its allocation between different sectors.

Another way of modelling the influence of social factors on human capital is the assumption that social factors directly affect the 'production technology' of human capital. In this case the constraint which is modified is not the rewards structure, but the human capital production function.

Several authors follow this approach. Cozzi (1998) for example, assumes that culture directly affects the efficiency of labour factor, and Iyigun et al. (2001) follow a similar argument. According to these authors, education acquired through formal schooling and ideology are two productive factors that are substitutes in the human capital accumulation function. The idea behind this assumption is that a more sophisticated ideology allows us to obtain more accurate knowledge about the facts, such that individuals with different ideologies but the same level of education make different inferences about the world.

In both models the production function of human capital can be represented as,

$$h_{i,t+1} = f(E_t; I_t) \tag{1}$$

where E_t is the education acquired in formal schooling and I_t is the ideology (or the culture) which can be ranked according to its degree of sophistication.[13]

Although these models contain a clear improvement with respect to the simple Beckerian model, given by the recognition of the role played by social factors in the individual formation process, they still present several shortcomings. First of all, they are based on poorly defined analytical

categories. A second problem, strictly linked to the first, is that it is not clear in what way ideology or culture is formed and how their formation process can be embedded in an individual choice-based model.

Another strand of literature which uses this analytical method involves attention to social interactions. In these models one typically postulates that human capital is formed by way of social relationships because these favour knowledge transmission among individuals (Benabou, 1996; Durlauf, 1996; Galor and Tsiddon, 1997; Hassler and Rodriquez Mora, 2000). Different cases of social knowledge transmission are identified: one of great importance is that occurring within the family because of the close and enduring relationships among its members. Yet also the relationships arising between members of the same group or community are important channels through which knowledge is transmitted. The models developed by Benabou (1996) and Durlauf (1996) attach greater importance to the latter channel of knowledge transmission.

In particular, Benabou's model highlights how articulate social relations can be and how their possible results in terms of human capital level and growth rate may depend on the forms which they take. He postulates that social interactions at community level may give rise to very different results in terms of human capital according to which form they take. If individuals with greater human capital exert considerable influence in the group (that is, the more educated members of the group are emulated), a community made up of heterogeneous individuals will be more efficient in terms of the human capital produced, because the 'high tails' of the distribution will prevail. Vice-versa, if the influence of those with a low human capital predominates, an increase in the proportion of high-quality individuals will have a negligible effect on human capital since the 'low tails' of the distribution will predominate. These different types of social interactions can be formalised by two production technologies of $h_{i,t+1}$. The first case is,

$$h_{i,t+1} = d_i(i,l)\varepsilon(\sigma)\bar{h}_{l,t} \tag{2a}$$

while the second case,

$$h_{i,t+1} = \frac{d_i(i,l)\bar{h}_{l,t}}{\varepsilon(\sigma)}, \tag{2b}$$

where $d_i\ (i, l)$ is the social distance function between individual i and l reference group, here considered as an exogenous parameter, $0 \le \sigma < +\infty$ is an index of the variability of the distribution of h over the group, $\varepsilon(\sigma)$ is an increasing function of this variability ($\varepsilon(0) = 1$), and $\bar{h}_l$ is the average level of human capital in l group.

Equation (2a) states that when the influence of more able individuals prevails, the transmitted human capital is above the average of the group, and the intensity of the transmission increases with an increase in variability of the distribution of *h* over the group. Vice-versa, equation (2b) states that if the influence of less able individuals prevails, the transmitted human capital is below the average, and an increase in variability further reduces this transmission.

Although neither this model analyses the formation of social interactions, it shows that their introduction helps to describe the formation of individual human capital in a clearer and more detailed manner by allowing us to distinguish cases where social relations have positive effects, from cases where their effects are negative. Moreover, through the social distance function it is possible to make social interactions endogenous. In fact, although in the Benabou model the social distance function is an exogenous parameter, the analysis could be extended by considering it as a choice variable which comes from a decisional process where an individual chooses the optimal social distance.

Another aspect greatly emphasised by this literature is that the effects of social interactions on human capital accumulation may explain the dynamic of the relation between inequality and growth.[14] In fact in the models of Benabou (1996) and Durlauf (1996), an equilibrium may emerge, where different groups of agents are formed. In this case individuals differ in their level and possibility of further accumulating human capital.

This result is particularly interesting since these groups can be considered as different social classes, each being characterised by a different level of human capital and a different possibility of further accumulating it. Such a class structure, however, is radically different from that of the capitalists–workers dichotomy that was prevalent in Europe in the 19th century and beginning of the 20th century, since in this case there would be no single class of workers, but a whole range of worker classes, each characterised by a different level of human capital and a different access to the sources of human capital production.[15]

The models of Galor and Tsiddon (1994 and 1997) and of Hassler and Rodriquez Mora (2000) mainly focused their attention on knowledge transmission occurring within a family. In Galor and Tsiddon's (1994) model, parental relations affect the level of human capital through two variables: the amount of parental knowledge and the resources invested by the individual in education, which depends on the parents' human capital.[16] Moreover, the authors assume that the parent/child relationship affects human capital formation not only by transmitting directly human capital to their offspring, but also by modifying their cognitive capacity. Formally, knowledge transmission at the family level can be written as:

$$h_{i,t+1} = h_{j,t} d_i(i,j)\alpha(I,h_{j,t})x_t(h_{j,t}) \qquad (3)$$

where x_t is the amount of resources invested in education, $d(i;j)$ is the social distance between parents and children, which is weighted by a factor $\alpha(I,h_{j,t})$. This factor is increasing in parents' human capital because individuals with parents who possess high *h* learn more quickly and are able to use their knowledge more efficiently, and depends on some family characteristics (I), such as the possibility for the parents to spend many hours with children, or whether the family structure is a 'nuclear' or 'extended' one. These institutional characteristics, as Coleman (1988) has emphasised, influence the quality of family relationships and hence the effectiveness of knowledge transmission.

In a more recent work, Galor and Tsiddon (1997) considered another feature affecting the transmission between parents and children. In this paper, they assume that knowledge transmission depends also on the occupations chosen by children. If they choose the same occupations as their parents, then human capital transmission is complete.[17] This happens because in performing a job, individuals develop a set of interpersonal relations that constitute a sort of 'social capital' to be used in their work, which can be transmitted to the children only if they remain in the same sector.

This case is particularly interesting since it shows that individuals are not members of a single group, but they are involved in different types of social relations, and it may be relevant to analyse how the relations developed in different contexts interact with each other. This idea has been well defined by Gellner (1996), who introduced the concept of 'modularity' of individuals, which means that individuals define themselves by multiple attributes associated with distinct spheres of social life. This implies that an individual may belong to different groups. In some cases there is no relation between the attributes an individual has in a group with those that s/he has in another, but in other cases this relation exists, and it may influence inter-group interactions, as in the case analysed by Galor and Tsiddon (1997).

9.3.2. Models with Social Factors which Modify Preferences

The analytical strategy followed by the models surveyed above, even if it is a straightforward way to include social factors in economic analysis, is actually rather limiting, since the decisional process which leads to the choice of a particular behaviour is not explicitly analysed.

A different way of modelling the growth effects of social factors is to assume that they modify the preferences structure or, more generally, the objective function. In this way the formation process of the social factor emerges as a solution of an individual decisional problem. However, even if this method may be more suitable to overcome the above cited problems, within the new growth theory it is not so widely used: the few papers which use it largely coincide with those that have focused on the demand for social status.

Generally it is assumed that social status depends on some individual traits, usually one's wealth or the level of education, and that an agent chooses the level of wealth or of education, in order to maximise his/her own utility function, where the latter is defined not only over a set of market goods, but also over the social prestige s/he obtains. However, social prestige can be defined only in relative terms. Consequently, to establish the relative social position of an individual, it is necessary to identify the reference group with respect to which the individual's social status is defined. Therefore, most of these models have to define 'a priori' the neighbour structure to which the individual refers.

Formally the individual choice problem is,

$$h_i^* \in \arg\max U_i\left(p, s(h_i, \bar{h}_l)\right) \tag{4}$$

where p is a vector of personal characteristics, including income, $\bar{h}_l$ is the average level of human capital in the reference group, s ($s_{h_i} > 0$ and $s_{\bar{h}_l} < 0$) is the social status function, implying that an individual obtains greater status if s/he differentiates her/himself from own reference group.

Although in this case the social factor (i.e. individual social status) is a choice variable whose level emerges as a solution of a maximisation problem, this framework is still not sufficiently general, since the formation of the reference group and, more generally, the structure of the social interaction environment do not emerge as an equilibrium solution.

Once again, there is a clear need to elaborate a more general framework that is able to explain how economic variables interact with the social environment, making the formation of the latter, at least partially, endogenous to the model. Moreover, it would be worth constructing a general framework which can unify the various models and the different analytical categories used to analyse the effects of social factors on economic performance.

9.4. FURTHER DEVELOPMENTS

Recently, there have been some interesting developments in two different strands of literature that may constitute some possible solutions to the above discussed problems. One is the model of individual social capital investment,

proposed by Glaeser et al. (2000); another can be found in recent papers belonging to the literature on non-market interactions. In what follows we will discuss the relevant features of these two approaches that, if embedded in a growth model, may lead to important developments for the analysis of the effects of social structure on economic performance.

9.4.1. The Individual Social Capital Investment Approach

In this model agents can accumulate individual social capital in the same way they do with human and physical capital. Individual social capital is defined as 'a person's social characteristics … which enable him to reap market and non market returns form interactions with others. As such, individual social capital might be seen as the social component of human capital' (Glaeser et al., 2000, p. 4). Moreover, this particular kind of investment is assumed to be time-consuming, and as such has an opportunity cost given by the hourly wage.

This definition enables the authors to use the standard model of optimal individual investment decisions to analyse the formation of individual social capital. By applying this very standard model, they are able to obtain interesting results: a relation between the lifecycle of an agent and the individual investment in social capital, which is positive in the early stages of life and negative in the latter stages; a negative relation between mobility and social capital investment, which implies that what reduces mobility, such as homeownership, also increases social capital investment; a positive relation between the patience rate across individuals and social capital investment, which generates a reduced form correlation between the latter and human capital accumulation.

An important feature of individual social capital, characterising also human and physical capital albeit to a lesser extent, is the presence of major externality and positive complementarity effects. Complementarities raise the possibility that there exist multiple equilibria in the levels of social capital investment, and explain how small differences in initial conditions may generate large divergences in long-run levels of social capital. Moreover, the presence of externalities and complementarities implies that the transition to the aggregate level is not immediate. Indeed aggregate social capital is defined as the average of individuals' social capital, adjusted for all the externalities, which are of a considerable amount and can be positive or negative. This latter feature depends on the type of individual social capital accumulated. For example, joining a network is a form of individual social capital that creates positive externalities, while status-seeking behaviour, which is another form, causes negative externalities to other agents in the reference group.

This approach has the great advantage of being able to give a unified interpretation for the different concepts used to analyse the effects of social factors on economic variables, using an analytical format very familiar to economists. In this framework, social interactions are only different forms of individual social capital: joining a group and status-seeking behaviour are only two of the many forms of individual social capital. Moreover, this model provides a coherent interpretation for the positive relationship between human capital and social capital: these are two distinct accumulable factors that show a positive relationship between them, because of an equal response to changes in the intertemporal preferences rate of individuals.

Nevertheless, by adopting this approach one may lose a major aspect highlighted by most of the papers analysed here: the fact that individual decisions are influenced by actions of other agents in a way that is not regulated (because it is not possible) by the market. By assuming that one can define a market price for the investment in social capital, it is difficult to maintain this specific nature of personal relations, which cannot be regulated by the market and, consequently, cannot be regulated by a price mechanism. Even if this aspect is treated through externalities, which express the direct effects of agents' actions on others, nevertheless, a direct analysis of these phenomena could be more correct, because the incompleteness of markets does not mean that an individual is unaware of the effects of such personal interactions. By confining them within the narrow space of generic externality effects, one makes this intuition unable to clearly emerge.

These considerations suggest that the most suitable framework for analysing the economic effects of social factors, that least distorts their real nature, may be the non-market interaction framework, where each person's actions change not only because of the change in fundamentals, but also because of the change in the behaviour of his/her own neighbours.

9.4.2. Non-market Interactions Approach

Another possible solution may be supplied by the literature on social interactions, which uses the random field approach, imported from statistical physics, where one typically postulates individual's interdependence and analyses the macrobehaviour that emerges. Follmer (1974) was the first paper to use this framework. Other models inspired by statistical physics are Scheinkman and Woodford (1994), that studies the impact of independent sectoral shocks on aggregate fluctuations, Glaeser et al. (1996) who use the voter model to analyse the distribution of crime across American cities, and Brock and Durlauf (2001), who develop a model which extends the random field approach to global interactions, to the case of discrete choices.

An interesting contribution is provided by Glaeser and Scheinkman (2000) who present a general model that is able to treat as special cases several of the better known models in this area. Hence, to clarify the main features of this approach, I will refer to the analytical format presented in this model.

In Glaeser and Scheinkman's model utility function includes the individual's actions, the actions of agents within the reference groups, individual's personal characteristics and common prices. The reference groups may include only individual's closest neighbour or the entire economy. Therefore this framework can examine both local and global interactions. Formally the individual problem is

$$\max\ U_i = U_i(a_i, A_i^1, \ldots\ldots, A_i^K, \theta_i, p) \qquad (5)$$

where

$$A_i^K = \sum_{j=1}^{n} \gamma_{i,j}^K a_j$$

($\gamma_{i,j}^K \geq 0$ and $\sum \gamma_{i,j}^K = 1$), $a_i \in I$ [18] is the agent *i* action, p is a vector of parameters and θ_i is a 'taste shock' of each agent. In other words the utility of an agent *i* depends on his own chosen action, on a weighted average of the actions chosen by agents in his/her reference groups (A_i^K), on his/her taste shock, and on a set of parameters.

An interesting result of this model is that multiple equilibria may arise even with very little heterogeneity. In this case two populations with slightly distinct realisations of the θ_i's could exhibit very different average values of the actions. This happens if the marginal utility of an agent's own action is more influenced by change in the average action of his/her peers than by a change on his/her own action. In other words multiple equilibria occur when the group effect is strong enough. An implication of this result is that it is possible to analyse how different groups of agents emerge as an equilibrium solution.

This model also has the potential to facilitate a more rigorous analysis of social capital. In fact, if social capital is interpreted according to the definition proposed by Durlauf (1999a), by which social capital is 'the influence which the behaviors of one's reference groups have on one's assessment of alternative courses of behavior' (Ibidem, p. 2), this can be identified, at least as a first approximation, as the weights in the A_i^K terms, because the latter terms capture the influence that the average behaviour has on the optimal choice of an agent. Moreover, as Brock and Durlauf (2001) have shown, by choosing the weights parameters appropriately, it is also possible to consider cases where agents wish to differentiate themselves from their own reference groups (status-seeking behaviour).

This approach could be very useful to highlight the sociality of an individual behaviour, providing, at the same time, a very general framework that includes as special cases all the possible ways in which social factors influence individual economic decisions. Moreover, if it is applied to continuous actions, it can explain how human capital formation depends strictly on social relations, emphasising that individuals are aware of such influences and they can act to modify their social interaction environment.

9.5. CONCLUSIONS

The literature analysed in the previous sections constitutes a new and interesting strand of research which seeks to integrate non-market mediated social interactions into the analysis of economic growth. There has clearly been an excessive proliferation of concepts and analytical categories, none of which are well defined, with all referring to the same phenomenon: the interdependence of agents not regulated by market mechanisms. In this paper I have set out to argue that by basing the analysis on the better defined concept of social interactions, it is possible to obtain a coherent and well-grounded analysis of the effects of social factors on economic growth. Consequently, an important stage will be to encompass in the endogenous growth framework models designed to analyse non-market interactions.

Finally, an interesting finding from this literature is that social interactions accentuate the heterogeneity of agents and may create distinct groups of agents. Each of these groups has its own specific rules of behaviour and different levels of transmittable human capital. This entails the emergence of a class structure based on the level and/or type of human capital possessed by individuals and on the possibility of further accumulation of such capital.

NOTES

1. See for example the F-connection theory of Ben-Porath (1980).
2. Where social classes are defined according to their function in the production process and the posses of an accumulating productive factor.
3. In this line of research are Galor and Moav (2000) who propose an evolutionary framework where some types of behaviours, such as preference for ‘quality’ children rather than for ‘quantity’ children, emerge as a result of an interplay between the technological progress and the fertility rate.
4. For example by the communication media, the school, etc.

5. This is a rather restrictive hypothesis, in fact, since humans do communicate about all manner of things. Nevertheless, the assumption proves very useful because it considerably simplifies the analysis (Manski, 2000).
6. Smith (1759, p. 3) wrote: ‘How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it’.
7. ‘Every man feels his own pleasures and his own pains more sensibly than those of other people … After himself, the members of his own family, those who usually live in the same house with him, his parents, his brothers and sisters, are naturally the objects of his warmest affection’ (Smith, 1759, p. 321).
8. More specifically, given a social space which comprises the structure of the agents’ neighbourhoods, the social distance between them is defined as ‘the number of links in the shortest path between the agent’ (Kirman, 1999, p. 24). According to this definition, those agents that are directly connected have a social distance which reaches the maximum value.
9. Max Weber (1930) developed the same idea, attributing the rise of capitalism and the Industrial Revolution to the Calvinist Reformation.
10. As we have already noted, Smith (1776) has recognised that individuals chose an occupation also for the reputation that they can acquire. Marshall (1890) also have noted that: ‘The desire to earn approval, or to avoid the contempt, of those around us is a stimulus to action … in any class of persons … A professional man ... will be very sensitive to the approval or disapproval of those in the same occupation’ (1890, reprinted 1962, p. 19).
11. See for a comprehensive survey Weiss and Fersthman (1998).
12. Carillo and Zazzaro (2001) have developed a neo-Schumpeterian growth model where social factors, such as the ‘professionalization process’ and the status seeking behaviour of professionals, modify the reward structure reducing the convenience to devote human capital to R&D sector and by slowing down the pace of technological innovation.
13. Iyigun et al. (2001) assume that ideology is in its turn affected by the level of education. This assumption combined with the possibility that ideological beliefs affect the human capital accumulation process generates a feedback loop between ideology and human capital.
14. In fact, this approach gives rise to what has been denominated as the *membership theory of inequality* (Durlauf, 1999b) according to which income distribution depends not only on individual characteristics, but also on characteristics of those groups to which an individual belongs.
15. It is interesting to note that this class structure seems to describe well what is observed in most OECD countries in the last part of 20th century since the end of seventies. In fact in almost all OECD countries it has been observed a strong increase in educational wage differential, with a spectacular increase in the return to education (Acemoglu, 1999; Aghion, Caroli and Garcia Pegnalosa, 1999; Goldin and Katz, 1999).
16. This happens because parents with high h are assumed to be willing to invest more in their children’s education since they have altruistic preferences and the welfare of their children is a normal good.
17. The same line of research has been pursued by Hassler and Rodriquez Mora (2000), who also emphasise that parental transmission is complete if there is no social mobility. The transmission of knowledge takes place between parents and children only within the entrepreneurial class and it is maximum if there is no innovation.
18. I can be a discrete set or an interval of a real line, therefore this framework can analyse discrete as well as continuous choices.

10. The evolutionary perspective on growth

Grazia D. Santangelo

10.1. INTRODUCTION

The interest in economic growth, central to the classical writers, re-emerged during the post-World War II period. The revised version of Solow's model (1957) has long been the most successful attempt in explaining total output growth. However, the widely recognised weakness of the model was the identification of technological change with the 'residual', which did not allow the importance of technological change to be specified. The main problem appeared to be the impossibility of distinguishing an increase in *per capita* capital due to a move along the production function from an increase due to a shift in this function.

Endogenisation of technological change has been the main aim of new growth theory (NGT) which, according to the positivist tradition of neoclassical economics, has proceeded by testing particular theories against data rather than describing and explaining phenomena as they are observed (Nelson, 1996). Therefore, although representing an advance by comparison to 'old' neoclassical growth models, NGT still proposes an understanding of economic growth as a smooth process based upon the concept of a continuing equilibrium, failing to encompass the historical account of technical change and related activities and institutions (Ibidem). By contrast, evolutionary theory adopts a behavioural approach to individual firms which, based on Simon's (1978) bounded rationality, perceives individuals as rational and maximising according to their limited mental processing abilities, information, and the complex environment in which they operate. This logical structure allows evolutionary theory to account for the macroeconomic phenomena explained by neoclassical theory by making more realistic assumptions, as well as investigate new research questions.

The aim of this paper is to provide a critical review of some selected evolutionary models which propose an alternative perspective to mainstream economic growth. It should be, however, underlined that no single

evolutionary model can account for all aggregate regularities at the same time, although a certain degree of consistency emerges between the different models (Dosi and Nelson, 1994). This is due to the fact that economic growth is a historical process, which interacts with, and is influenced by, many factors outside the economic domain (e.g. culture, institutions, etc.) (Verspagen, 2001). Moreover, due to the microeconomic perspective of economic change adopted by evolutionary theory as well as the concern with *ex post* explanation rather than *ex ante* prediction, all these models are rather different in terms both of analytical structure and issues analysed. Therefore, this does not intend to be an exhaustive review of evolutionary growth modelling. Nonetheless, we believe that the models herein discussed, selected on the basis of classification made by Dosi and Nelson (1994), provide a broad idea of the main features of the evolutionary perspective on growth. The common denominator of all the models reviewed here can be identified in the explicit firm-level microfoundation.[1]

The paper is organised in 6 main sections. The next section provides an overview of evolutionary theory as an alternative to orthodoxy in economics as presented in Nelson and Winter's seminal book (1982). The aim is to provide a background for the following sections as well as an understanding of the historical emergence of the evolutionary perspective. In this sense, it should be underlined that, after the 1982 watermark, evolutionary theory has developed its own research agenda without benchmarking against orthodox results. Section 10.3 reviews the basic evolutionary growth model. Further developments of evolutionary growth theory are covered in Section 10.4. Section 10.5 discusses the contributions of the evolutionary perspective to growth theory. A few conclusions are drawn in Section 10.6.

10.2. AN OVERVIEW OF EVOLUTIONARY (*VERSUS* ORTHODOX) THEORY

The emergence of evolutionary theory in the economic realm may be framed within the widespread dissatisfaction with the state of economic theory. This sense of malaise was mainly directed towards the inability of orthodox theory[2] to come to grips with empirical reality. In this context, the evolutionary critique of the mainstream economic approach has addressed three main fundamentals of the orthodox structural logic (Nelson and Winter, 1982). First, the continued reliance on equilibrium analysis (even in its more flexible forms) neglects phenomena associated with historical change, as disequilibrium situations are outside the focus of investigation. Second, the perfect-information assumption (even in its more relaxed forms) limits the explanatory power of the theory when facing complex realities by implying that economic agents can foresee all possible contingencies and weigh their consequences. Third, the economic rationality of the agent – in the sense that he/she optimises – encompasses clear anticipation, calculation and clarity of risks taken by this actor in dealing with realistic complexity without accounting for confusion about the situation, distraction and mistakes on the part of the actor.

Evolutionary theory criticises neoclassical theory, as the latter does not offer an explanation of why decisional rules are what they are due to its two structural pillars: maximising behaviour and the concept of equilibrium. The rejection of the notion of maximising behaviour can be traced back to the consequent rejection of the distinction made by orthodox theory between the choice set and choosing. In fact, according to the evolutionary viewpoint, the economic agent engaging in the decision-making process is not aware of all possible alternatives he/she can have, but (some of) those alternatives will be discovered during this process by trial and error. In this sense, the determination of decisions and the decisional outcome involve a stochastic element and may be non-optimal. Therefore, agents' behaviour cannot be assumed to be prompt and rational. As far as the concept of equilibrium is concerned, although the analysis of a stable equilibrium configuration in which particular forces no longer produce change is recognised as important, the drastic narrowing of orthodox investigation to the equilibrium situation is firmly rejected as it allows explanation of the relations among the efficient survivors, but not the competitive struggle consumed in disequilibria. Conversely, the latter aspect is captured in the evolutionary perspective by the concept of selection.

Going into the details of evolutionary economics, this is built upon population thinking rather than typological thinking as in the orthodox approach. Indeed, the concern is with frequencies of behaviours that differ, not with uniform behaviour. While in the population perspective variety is a natural state, variations around the ideal type are regarded as accidental. Each firm has its interpretation of economic opportunities and constraints and, consequently, different firms perform different actions in different ways. This is due to the fact that the firm's capabilities are embedded in its organisational structure, which allows it to pursue certain strategies more easily than others. Unlike the orthodox theory, in a highly uncertain environment there is no global optimisation over a given set of choices comprising all objectively available alternatives. Drawing on the managerial school, evolutionary theory goes on to claim that firms as such may not have any objective in the sense that what is required to operate in the business world is just a procedure for determining the action to be taken, rather than a defined object function. In evolutionary thinking, this procedure depends upon the concept of 'routines' which can be understood as genes in

evolutionary biological theory. Routines, defined as 'a repetitive pattern of activity in an entire organisation' (Nelson and Winter, 1982, 97), are persistent features of the firm as well as determinants (together with the external environment) of the firm's behaviour. This implies that choice-making is a firm-specific process. Like genes, routines are heritable as tomorrow's plants generated by today's firms have the same characteristics (*path-dependency*). They are selectable as shown by the fact that firms owning certain routines perform better and increase their importance over time through a *continuous selection*. This selection mechanism is well described by the biological concept of natural selection and evolution, hence the label 'evolutionary' applied to this new economic approach.[3] In this sense, evolutionary theory is concerned with a dynamic rather than static analysis. In this process, firms with different routines compete in a rival selective process by attempting to differentiate one from another (*variety*). Firms' operations are framed within an uncertain environment by which they are affected (*mutation*) and to which they adapt by trial and error (*adaptation*). In this sense, the environment pushes firms into continuous learning and adaptation (both of which are firm-specific processes). Competition is then understood in terms of dynamic selection over time promoting economic change because new techniques progressively displace outmoded ones. This change proceeds gradually and in a *localised* manner as exploration and development of new techniques is likely to occur in the neighbourhood of the techniques already in use, as already highlighted by Atkinson and Stiglitz (1969) in their critique of the neoclassical concept of technological progress.

Drawing on the Schumpeterian tradition, in evolutionary theory change arises partly from within the firm and part as the selective effect of the environment. Due to the struggle for survival engaged by firms, technology is a crucial factor in the selection process. Successful firms are major technological innovators and/or imitators[4] as in the capitalist reality technology-based competitiveness is a source of major advantages (Schumpeter, 1942). This may be better understood when recalling the definition of technology provided by evolutionary theory. Unlike the orthodox tradition, evolutionists distinguish between a *public* and a *tacit* element of knowledge (Nelson and Winter, 1977, 1982). The former, better defined as potentially public knowledge, can be codified in patents, blueprint, textbooks, etc., implying that 'public' technology can be easily transferred and, therefore, traded between firms – this corresponds to the neoclassical understanding of technology. The latter is a private element embodied in the firm's organisational routines, expertise and skills acquired through a process of learning, and takes the shape of a firm-specific set of practices. If the public aspect of technology can be easily transferred, the tacit element is non-tradable and difficult to transfer as the result of a process of learning-by-using (Rosenberg, 1982). Therefore, since tacit knowledge is developed through an internal learning process, technology is also partially *context-specific* in the sense that it develops linkages with the local environment.[5] This aspect is strictly linked to the concept of the technological paradigm[6] as well as the concept of institutions (understood in both a narrow – e.g. governments, regulatory bodies, etc. – and broad sense – e.g. universities, business culture, tacit patterns of behaviour, etc.).

10.3. EVOLUTIONARY GROWTH THEORY

Drawing on the features of evolutionary theory outlined in Section 10.2, an evolutionary model is characterised by *dynamic analysis* accounting for *random* elements that generate some variation in the variable in question (e.g. the firm or the technology) through a *mechanism selecting on extant variation* (Dosi and Nelson, 1994). In this process, learning and discovery by trial and error play a great role in defining adaptation and mutation in the variable in question. Based on Schumpeter's (1942) idea of the centrality of the firm in capitalist economic development, evolutionary growth models are *microfounded* in the sense that firms are the key actors in making investments to develop and adopt new technologies, and in the use of these technologies to produce goods and services.

10.3.1. The Basic Model

The basic evolutionary growth model of Nelson and Winter (1982), developed in different stages (Nelson and Winter, 1974, Nelson, *et al.*, 1976), is based upon the logical apparatus set forth in Winter (1971) where decisions, rules, search and selection mechanisms are provided analytically.

It is assumed that each firm has the same convex constant returns to scale technology represented by a production set Y in 3-dimensional Euclidean space where it is assumed that there are two inputs and one homogeneous product. The generic element of Y is indicated by y. That is, each firm has the same technology defined by a $3\times M$ matrix A (where M is the number of techniques). Each column is a list of inputs and an output flow that is feasible in a single plant in a single period, and two rows are strictly negative. Moreover, in each period, a firm is characterised by a stock of capital K_j, which is assumed to be entirely employed. This structure provides the basis for the mechanism of the decision rule (for production decisions). This rule is represented simply by one column of A and can be formalised for each firm (j) in time t by the vector $r_{jt} = (1, a_{Lt}, a_{Kt})^T$, where a_{Lt} (a_{Kt}) indicates the amount

of labour (capital) needed to produce one unit of output. Thus, at any point in time, a state of a firm is defined as a pair of inputs a_{Lt}, a_{Kt} and its capital stock K_j. An industry state is simply a list of firms' states $[(r_{1}.,K_{1}.), \ldots, (r_{P}., K_{P}.)]$ (where P is the finite number of firms potentially in existence). An industry state in a given time period implies a certain aggregate capital stock $\sum K_j$, a certain aggregate output in use $\sum K_j / a_{Kj}$ and a certain aggregate labour demand $\sum a_{Lj} / a_{Kj}$.

The prices of the output and inputs are determined by the industry state according to a continuous demand price function, which maps the aggregate capital, labour and output quantities into a list of non-negative prices of output, capital and labour which ensure market equilibrium. It is also assumed that this function ensures positive profits for all firms for some industry states. It is assumed that the industry wage rate is determined according to the following rule:

$$w = a + b\left(\frac{L_t}{(1+g)^t}\right)^c \qquad (1)$$

where t is the time period, L_t is the aggregate demand of labour in the period and a, b, c and g are constants. If $g = 0$, labour supply conditions are constants over time and the model as a whole is a Markov process with constant transition probabilities. If $g \neq 0$, then labour supply conditions undergo changes and the model is still a Markov process but with time-dependent transition probabilities. In turn, the wage rate allows the firm's gross profitability to be determined. Gross investment is determined by gross profit.

In this context, natural selection is formalised by a finite Markov chain. From one period to the next, the state of an individual firm changes according to probabilistic rules that depend on the initial state of the firms and its profitability (determined by the initial state, the wage rate and constant parameters). Since the industry state is the list of firms' states, the transition probability rules for individual firms define the transition probability in the set of industry states. It is assumed that there are extant firms and potential entrants. As far as existing firms are concerned, transition probability rules for productive techniques are specified on the grounds of firms' gross return on capital according to whether the firms belong to group 1 or group 2, defined as follows:

GROUP 1: firms with positive capital in the current state satisfice with respect to their decision rules if they are sufficiently profitable – that is if they make a gross return on capital exceeding a target level. Therefore, they retain the production technique of the state with probability 1 (*satisficing*);

GROUP 2: firms that make a gross return on their capital less than the target level undergo a probabilistic technique-change process. This process may entail either seeking incremental improvements in the firm's present methods (*local search*) or looking at what other firms are doing (*imitation*), but not both at the same time.[7]

(a) In the *local search* case, the search is local in the sense that the probability distribution of what is found is concentrated on techniques close to the current one. Distance between technique *h* and *h'* is given by

$$D\,(h, h') = \mathrm{WTL}\,|\log a_L^{h} - \log a_L^{h'}| + \mathrm{WTK}\,|\log a_K^{h} - \log a_K^{h'}| \qquad (2)$$

where WTL+WTK = 1. That is, the distance between the two techniques is a weighted average of the absolute differences in the logs of input coefficients. The closer the techniques to each other, the higher the probabilities for transition from one to another.

(b) In the case of *imitation*, the probability that firms looking at what other firms are doing will find a particular technique is proportional to the fraction of total industry output produced by that technique in the period in question. Alternative techniques turned up in the search process are adopted by the firm only if they yield a higher return per unit of capital than the firm's current rule.

As far as entrance is concerned, potential entrants are classified into two groups:

GROUP 1: firms with zero capital in the current state contemplating the use of a production decision rule that implies a return to capital exceeding the target level. These firms become actual entrants with a given size probability less than one;

GROUP 2: firms contemplating rules that yield capital return less than the target level and do not enter with probability 1.

By means of computer simulation, the authors used this analytical framework to show that an evolutionary model of the sort described above is able to account for the macroeconomic patterns explained by neoclassical theory, as represented by Solow's 1957 article, where data on gross national product, capital input, labour input and factor prices over forty years are considered. Nelson and Winter (1982) generates the same macro aggregates by building them up from microeconomic data (e.g. time paths of firms' and industry

inputs and outputs, time path of the industry wage rate and firms' and industry rate of return on capital). Setting the original conditions of the model in order to make them correspond to the conditions revealed in Solow's data, the two authors obtain a smoother aggregated 'technical progress' than that found by Solow in the real data for the US, thereby confirming the gradual and incremental character of technical change hypothesised by evolutionary theory. Moreover, the simulation model does generate 'technical progress' with rising output per worker, a rising wage rate and a rising capital–labour ratio, and a roughly constant rate of return on capital.

However, it is worth emphasising that, although in the simulation model firms respond to profitability signals in making technical changes and investment decisions, they are not maximising profit. Conversely, emphasis is placed on corporate behaviour: firms, that are doing well, relax by making only small changes when they do change their decision rules; firms subject to payout constraints grow by maximising investments. Similarly, the model does not rely on the concept of equilibrium. At any point in time, there are different techniques used and rates of return obtained as a result of the fact that there are always better techniques not being used. It assumes the existence of a set of physically possible techniques, a subset of which is not assumed to be known at each particular time. Rather, this subset is explored historically through an incremental process where non-market information flows among firms play a major role and firms know only one technique at the time.[8]

10.4. FURTHER DEVELOPMENTS: THE ROLE OF TECHNOLOGY DIFFUSION

The Nelson and Winter (1982) model has been further developed, giving rise to two streams of evolutionary modelling attempting to tie together analysis of diffusion patterns and productivity change explaining technological asymmetries (i.e. gaps between firms in terms of costs of production and product characteristics), technological variety (i.e. diversity related to differences between firms in their search procedures, input combinations and products, despite similarity in their production costs) and behavioural diversity (i.e. within the same industry firms show different strategies). The issues addressed in these models concern the processes involved in economic development. Although development issues lie outside the focus of this survey, there is scope for including these models as their results (e.g. the significance of the learning process internal and external to the firm, of the proprietary aspect of technology, and of embodied technological change) can provide significant insights in explaining growth.

In discussing these models, a distinction is made between models allowing two technologies (an old and a new one) and models allowing a multitude of technologies. Common to both models is the elimination of the stochastic element of the new-technique generation focusing on a selection of techniques that are initially in use. In the discussion, emphasis is placed on the main results obtained, while providing a brief intuitive overview of the model.

10.4.1. Diffusion in a Two-technology Model

In the model proposed by Silverberg, Dosi and Orsenigo (1988), industry level demand is taken as given and growing at some exponential rate. Firms own some market shares of this demand at any given time, although market shares may change over time with a characteristic time constant as a dynamic response to disparities in relative competitiveness. For each firm, competitiveness is defined as a linear combination of market price and current delivery delay. Firms' market share changes according to the deviation of the firms' competitiveness from the industrial average competitiveness. It is assumed that entrepreneurs are aware of the process of economic growth and technological change, whose developments are taken into account when deciding on their fixed investments. Thus, the decision-making process is based on certain rules of thumb and animal spirits in the form of decisional rules governing replacement policy and expansion capacity. Technical change is embodied in vintages. The capital stock of each firm is represented as an aggregate of gross investment made between the current period and the scrapping date. This capital stock may consist of different technologies as well as different vintages of a single technology trajectory. The change of capital stock over time is defined as a linear combination of net expansion, gross investments and removals due to scrapping. The desired level of capacity expansion may be set initially at any level, but revised over time in order to account for the deviation of the rate of capacity utilisation from its desired level. Labour is the only current production cost firms face. The total quantity of labour per unit of capital is given by the average of the technical labour-output coefficient at different times comprised between the current period and the scrapping date weighted by the vintage. Changes in the total quantity of labour over time are due to more productive equipment through investments and removal of equipment due to scrapping. Changes in the levels of production reflect deviation of the current delivery date from an industry-wide standard level. Prices are

determined as a dynamic compromise between the desired mark-up on unit costs and relative competitiveness.

The core of the model concerns the comparison of two technological trajectories representing at any time the maximum productivities achievable in best practice vintages of respective technologies. It is assumed that these technologies change at some rate and that the second is always superior in productivity. Initially, all firms use technology 1, which is already mature in the sense that skills levels are saturated at 100 per cent. As far as technology 2 is concerned, firms possess lower efficiency and are unaware of the margin for further development. Therefore, they can only guess at the rate of further improvements in efficiency and technological progress. This reflects the fact that the productivity of a technology concerns specific expertise and experience internal and external to the firm rather than the machines used. Therefore, investment decisions rely on the ability of evaluating the prospects for further development either by acquiring experience with the best practice technology or by waiting for the right moment in order to avoid possible development costs. The rate of change of the internal skills level of firm *i* using the new technology, s_i, follows the rule

$$s_i = \frac{AP_i}{CP_i + C} s_i(1 - s_i) \qquad \text{if } s_i > s_p \qquad (3)$$

where A is a parameter, P_i is the firm's current production, CP_i is its cumulated production with the new technology, C is a constant proportional to corporate capital stock and s_p is the level of skill generally available in the industry also to those firms which are not yet producing on the new technological trajectory. This 'public skill' can be thought of as skilled labour and management moving between firms, operating instructions of some industrial equipment diffused by manufacturers to users, industrial and trade publications, etc. Thus, the rate at which the level of generally available skill changes can be formalised as

$$\dot{s}_p = A(\langle s \rangle - s_p) \qquad (4)$$

where $\langle s \rangle = \sum_i^N f_i s_i$ with constant parameter f_i and N is the number of firms in the industry. Firms enjoy this learning externality even if they are not yet employing the new technology: $\dot{s}_i = \dot{s}_p$ if $s_i = s_p$. In deciding whether to switch to the new technology firms may abandon their investment criteria indicated above in order to take into consideration the gains in productivity due to the installation of the new equipment and to the early proficiency in use. Such considerations depend on how optimistic firms are about the future development potential of the new technological trajectory. Thus, firms select an 'anticipation bonus' (X_i) they award to the new technology in making their choice of technique. The new technology will be preferred to the old according to the following rule:

$$\frac{(P_2 - P_1) s_i X_i}{c_1 - c_2} \leq b_i \tag{5}$$

where P_1 and P_2 are the prices of the old and new technique, respectively; c_1 and c_2 are the respective unit costs of the old and new technique and b_i is the target payback period of firm *i*. Therefore, the new technique will be preferred to the old if its adjusted productivity is higher and if it is cheaper ($P_2 < P_1$). If it is more expensive ($P_2 > P_1$) it will still be preferred provided the difference in price is counterbalanced by lower unit costs ($c_2 < c_1$) within the desired payback period.

Due to the complex mathematical properties of the model, the authors run computer simulations in order to uncover some of its economic properties. In doing so, they focus on the anticipation bonus reflecting expectations about the future course of the new trajectory in order to explore the strategic aspects of the diffusion process and the problem of the interdependence of behaviour. If all firms anticipate incurring the development costs associated with bringing the new technology to commercial maturity, none of them will adopt the new technology. Conversely, if the rate of internal learning is accelerated, this will raise the dynamic appropriability of the innovation of the early adopters, who show high efficiency levels. Equally the evolution of firm-specific and external skill levels allows middle and later adopters to start from higher initial levels of efficiency due to external learning, and to overtake the early adopters. Therefore, technological innovation and diffusion are characterised by collective effects and social tensions between private and social gain, which impact on the aggregate economic growth. The rate of internal and external learning affects firms' investment decisions concerning production techniques. In turn, this affects corporate productivity and growth, which impact on aggregate economic growth.

10.4.2. Diffusion in a Multiple-technology Model

In Soete and Turner (1984), it is assumed that there is a large number of firms producing a homogeneous good by using different *N* techniques. The number of different techniques may change over time as a result of new technological innovations. Each technique uses two factors of production (capital and labour). For each technique, the capital–output ratio and labour–capital ratio may change over time due to an increase in efficiency. The capital stock of the economy is given by the sum of the capital stocks used

by each technique. No depreciation or scrapping is assumed. Assuming a common wage rate, the rate of return on capital obtained by all firms using the α technique is denoted by r_α and the return on capital for the whole economy is denoted by *r*. A particular technique is the best practise technique if it yields the highest rate of return given the common wage rate. The entrepreneurs will search for the most profitable technique to invest in, but not all will be successful in that search. The underlying behavioural assumption is that entrepreneurs make different decisions as a result of 'retardation' factors such as uncertainty about the profitability of the new technique, costs and time involved in learning and using it, possible protection, etc. Therefore, the adoption of the new technique may be delayed until information about other firms' experience is available. This behavioural pattern, which is reflected in the entrepreneurs looking for greater profitability, who will not adopt the most profitable technique immediately because of some retardation factors, can be formalised in the following investment function for technique α:

$$I_\alpha = \sigma r_\alpha K_\alpha + \eta\, (r_\alpha - r)\, K_\alpha \tag{6}$$

where $\sigma < 1$ is the saving ratio out of profits assumed to be the same for all techniques, K_α is the total capital stock embodied in technique type α and η is a constant. If $\eta = 0$, investments in technique α will come only from the surplus generated by the α technique ($\sigma r_\alpha K_\alpha$); if $\eta \neq 0$, then investment in technique α will be greater or less than the surplus generated by the α technique depending on whether $r_\alpha > r$. That is, a more profitable technique will attract investment which, conversely, will decline if the technique is not very profitable. Given $g = \sigma r$ and $g_\alpha = (\sigma + \eta)\, r_\alpha - \eta r$, the differential growth rate of each technique ($g_\alpha - g$) is given by

$$g_\alpha - g = (\sigma + \eta)\, (r_\alpha - r) \tag{7}$$

This implies that the difference in the growth rate of a particular technique α is proportional to the difference between the profitability of this technique and the profitability of the whole economy.

The authors show that the rate of technological change, v, is the following

$$v = \sum_a \left.\frac{\partial r\alpha}{\partial t}\right|_w \frac{K_a}{K} + (\sigma + \eta) \sum_\alpha (r_\alpha - r)^2 \frac{K_\alpha}{K} \tag{8}$$

Two components can be distinguished: the first provides the weighted average of the rate of technical progress for each technique α and can be referred to as disembodied technical change (i.e. accounting for the mere occurrence of innovation); the second, which represents the 'diffusion' term or technological change embodied in new investments, gives us the covariance of the deviation of the rate of return with respect to the distribution K_α/K. If there is a wide distribution of the rates of return of different techniques, the contribution of the diffusion terms to the overall rate of technical change will be high. Indeed, in this case there will be a large number of entrepreneurs spread across the spectrum of rates of return. Thus, the number of entrepreneurs moving to new techniques will be high. This implies a rapid diffusion and a high rate of embodied technological progress. Conversely, a sharp distribution of the rates of return of different techniques will reduce the contribution of the diffusion term to the overall rate of technical change, as the number of entrepreneurs moving to the new technique will be small. Therefore, aggregate economic growth is related to variation across technologies and their levels of diffusion. In turn, this implies that technological progress embodied in new investments increases the diffusion process of new techniques and accelerates aggregate growth.

10.5. EVOLUTIONARY CONTRIBUTIONS TO GROWTH THEORY

In what follows, evaluation of the overall contribution of evolutionary growth theory is attempted beyond the models discussed in the previous pages.

According to Nelson and Winter (1982), the development of traditional growth theory along unsatisfactory lines may be attributed to the detachment between formal and appreciative theory (Ibidem). The latter tends to be close to empirical work and concerns what the analyst thinks is going on. Therefore, it can be usually expressed in terms of storytelling. The former appeals to data for support, as its main purpose is to set up an abstract structure enabling proposed logical connections to be explored, found and checked (Nelson and Winter, 1982; Nelson, 1998). If economic theorising is going well, then empirical facts influence appreciative theory, and appreciative theory provides the basis for formal theory. However, according to the above authors, in the history of economic thought what seems to have happened is that formal theory and appreciative theory developed separately, with formal theory sourcing aspects of appreciative theory which could have been straightforwardly formalised according to the available mathematical tools and leaving the rest aside due to the lack of analytical tools to model non-competitive general equilibria.[9] This detachment marked a twofold development of the analysis of economic growth: an analysis of economic growth at macro level concerned with formalisation and an analysis of economic growth at micro level which is more concerned with empirical evidence.

On the basis of the methodological distinction between formal and appreciative theory and under more realistic assumptions, the evolutionary perspective has been able to account for the facts explained by neoclassical theory as well as compile its own research agenda, thus providing some major contributions to growth theory. First of all, due to its microfounded nature evolutionary theory treats the micro processes as fundamental and the macro processes as aggregates. This implies that macro phenomena are the results of underlying micro phenomena, which play a major role in explaining aggregated processes. The great emphasis placed on these micro phenomena has allowed evolutionary theory to deal with structural change issues more realistically. The adoption of an aggregated (Cobb-Douglas) production function – whose theoretical use is widely known to be logically faulty (e.g. see Harcourt, 1972) – prevents neoclassical as well as new growth theory from understanding the effects of important distributive changes that affect overall growth rates. The key difficulty lies in the fact that the overall growth rate of aggregate output and labour productivity depends on the cross-sectoral variations in the rates of output growth and level of labour productivity (Cornwall and Cornwall, 1994). By treating the economy as a homogeneous whole, new growth theory assumes that each sector experiences the same changes in productivity growth. This fails to capture the fact that rising productivity growth in only one sector impacts on average productivity growth according to the output share of the affected sector and any resulting inter-sectoral labour reallocation. Evolutionary analysis overcomes the limits of the orthodox theory, conducting a microeconomic analysis through the use of an evolving Leontief technology (as in Nelson and Winter’s model discussed in Section 10.3).[10]

Along these lines, the evolutionary perspective has contributed to more realistic theorising on growth by adopting a behavioural approach in explaining the corporate decision-making process. In this perspective, firms take their decisions on the grounds of some observable rules of thumb, which, taken as given, link environmental stimuli with corporate responses. This allows the construction of a more realistic growth theory than the neoclassical and new growth theory, where the rules concerning the firm’s decision-making process are deduced from maximisation. Adoption of a behavioural approach also enhances the explanatory power of evolutionary growth theory when considering the highly uncertain environment in which the firm operates. When taking their decisions, firms know nothing about the possible alternatives and potential outcomes of success. Therefore, it is impossible for them to face ‘large, well defined production sets that extend beyond the experiences range of operation’ (Nelson and Winter, 1982). Conversely, firms proceed by trials and errors through search processes.

A further contribution of evolutionary theory lies in the introduction of a dynamic concept of competition understood as evolving competitive advantage, which does not rely on equilibrium analysis as in the orthodox theory. The latter understands competition as circumstances where no relative competitive shifts or profits can be realised and assumes that the system must be near or in this state. Instead, in the evolutionary perspective competitions is defined in terms of conditions which continually change in response to the strategies pursued by firms and feedback from the rest of the system.

In this context, recognition of the proprietary aspects of technological change (against the traditional conception of knowledge as a public good) has been identified has a major element in understanding macroeconomic growth. The possibility of appropriating technological innovation generates differences in growth rates between firms and at aggregate level. The recognition by NGT of market imperfections, externalities from R&D investments or from education (Lucas, 1988), spillovers generating economies of scale (Romer, 1986) and of the proprietary aspect of technology – due to the profitability of R&D investments enabling the firm to appropriate a portion of the increase in productivity – can be seen as a consequence of heterodox developments. Similarly, the rise of an evolutionary alternative explanatory scheme has also promoted the flourishing of neo-Schumpeterian models within the NGT realm as in the case of Aghion and Howitt's 1992 model, which treats technological change in terms of a process of 'creative destruction'. Therefore, the main differences between these and the neoclassical models lie in more realistic assumptions of the former over the latter. Nonetheless, recognition of these phenomena as important determinants of growth cannot be regarded as a novelty given that applied theory has long emphasised such aspects. Moreover, as noted by Aghion and Howitt (1994), new growth theory models still have some limitations owing to their reliance upon rational expectations – the assumption of perfect information is relaxed only by treating uncertainty about the future in terms of a correctly specified probability distribution of possible future events – and the lack of attention to institutions and transaction costs. NGT models have taken into account the uncertainty that surrounds technological change, but without recognising that continuing technological advance implies a continuing state of disequilibrium. This is mainly due to the optimisation nature of the economic agents in the models. Therefore, although technical advance is regarded as the main source of economic growth, its concept is understood as involving shifting equilibria, whose paths are foreseen by the actors involved.

However, due to the evolutionary research agenda, which encompasses *formal* theory as a further subsequent abstraction of *appreciative* theory, attempts made to formalise the links between the micro and macro aspects of growth according to an evolutionary logic are rather recent. Nonetheless, there have been some significant attempts to link explicitly macro and micro aspects (for a review see e.g. Silverberg and Verspagen, 1998). For instance, along these lines a more recent contribution has been provided by Iwai (2000), who has shown that as long as the state of technology retains a feature of disequilibrium, the economy will keep generating positive profits and, in turn, economic disequilibrium due to continuous innovation. Moreover, some macro results achieved by NGT have been accounted for by evolutionary growth theory. However, viewing evolutionary growth theory as just able to explain phenomena explained by mainstream theory but with a more realistic microfoundation is a 'minimalist' position: evolutionary growth theory should be seen as an approach extending along its own lines and focusing on its specific methodological features (Silverberg and Verspagen, 1998). Indeed, after Nelson and Winter's 1982 formalisation, evolutionary (growth) models have developed along such lines (e.g. by attempting to endogenise the mutation and imitation process), rather than providing a benchmark against neoclassical results.

10.6. CONCLUSIONS

This paper reviews selected evolutionary models which propose an alternative view to mainstream economic growth theory. The models discussed can be framed in a wider attempt to build up an economic theory capable of explaining macroeconomic patterns under more realistic microeconomic assumptions than those made by orthodox theory. A major element of these models is the explicit assumption of bounded rationality, as actors behave according to their routine and know nothing about what is optimal. Similarly, the evolutionary character of such models lies in the fact that firms select technologies by deciding which to introduce as well as by deciding which one takes on board. In turn, on the basis of their fitness firms are selected in the market. Based on a theory of firms' technological change, evolutionary growth models seem to unfold the immediate source of growth by providing some insights into the nature of technology, the processes driving technological change and the factors influencing corporate behaviour and effectiveness.

The microfounded character of such models has allowed evolutionary theory to avoid the criticism levelled at neoclassical and NGT analysis based on macro aggregates. Moreover, the adoption of an approach based on observable patterns of behaviour rather than *a priori* theoretical assumptions as well as the proprietary aspect of technology makes the evolutionary setting more realistic, thus enhancing its explanatory power. Even if these features are specific to an evolutionary approach, some have been incorporated into NGT given their relevance to the understanding of economic growth. Similarly, realisation of technological change as a ‘creative destruction’ process has given rise to a neo-Schumpeterian stream of models developed within NGT. Nonetheless, although the neo-Schumpeterian endogenous models have partially attempted to account for the limits of orthodox growth theory, they still rely heavily on traditional neoclassical assumptions.

NOTES

1. It should be, however, pointed out that other models adopting a higher level of aggregation do exist (for a review see e.g. Silverberg and Verspagen, 1998).
2. Following Nelson and Winter (1982) ‘the orthodoxy … represents a modern formalisation and interpretation of the broader tradition of Western economic thought whose lines can be traced from Smith to Ricardo through Mill, Marshall, and Walras. Further, it is a theoretical orthodoxy, concerned directly with the methods of economic analysis and only indirectly with any specific questions of substance’.
3. It should be emphasised that in the literature (Foster, 1996) it has been remarked that biological theory provides a general model that explains a process of endogenous change by the interaction of several mechanisms (e.g. endogenous change, selection retention).. Thus spurious analogies between biological and economic phenomena are rejected.
4. As broadly discussed by Cantwell (1999), imitation also presumes firm-specific capabilities.
5. Technological advantage proceeds autonomously in terms of knowledge and inventions, but ‘innovation – i.e. application and diffusion of specific techniques in the productive sphere – is very much determined by social conditions and economic profit decisions’ (Perez, 1985).
6. In a pioneering article, Dosi (1982) defines a technological paradigm as ‘a model of solution of selected technological problems based on selected principles derived from the natural science and on selected material technologies’ (Ibidem, p. 152). Each paradigm shapes and constrains the rates and direction of technological change regardless of market inducements.
7. These means through which firms could generate novelty have been further expanded in other evolutionary ‘micro-models’ (for a review see e.g. Silverberg and Verspagen, 1998)
8. This leads to the evolutionary critique of growth accounting attempting to single out the relative contribution of different factors on growth given the impossibility of exactly specifying the production set (Nelson, 1973, 1996).
9. In fact, while in the 1950s and 1960s neoclassical theory was concerned with bringing together bodies of formal theory into a specific growth context and with formalising appreciative ideas in a naïf form, appreciative theory embarked on an analysis of micro-aspects of technology whose results started questioning some of the fundamentals of neoclassical economics (Nelson and Winter, 1982, Nelson, 1994).
10. The knife-edge problems of the Leontief production function affecting the Harrod–Domar model are overcome by the Leontief technology by introducing flexibility through innovation (Andersen, 2001). In fact, within this context disequilibrium on the labour market yields an active search by firm towards new methods of production which employ the most abundant factors of production more intensively.

11. Competition, rent seeking and growth: Smith versus the endogenous growth theory

Antonio D’Agata*

11.1. INTRODUCTION

Following Solow’s interpretation, growth theory is devoted to explaining the growth of *potential* capacity of economies, not their *effective* growth (see Solow, 1999, p. 639). The endogenous growth theory (EGT), however, has been able to explain not only sustained growth of potential capacities of economies, but, by making it clear that economic policies can affect the rate of growth, it has also been able to provide insights into the causes of different actual rates of growth among countries (see Rebelo, 1991).

As regards the explanation of the growth of potential capacity, the EGT has obtained persistent economic growth by removing the assumption of decreasing returns on accumulated factors and by assuming that savings are completely transformed into investment. These two features allowed Kurz and Salvadori (1998b) and, among others, D’Agata and Freni (this volume, ch. 2) to claim that the EGT view of the mechanics of growth can be traced back to the classical economists. As for the explanation of actual growth capacity, one of the main conclusions is that the actual growth of economies depends crucially upon the rewards earned by entrepreneurial resources through the free market. In fact, as entrepreneurs are not interested in profits in themselves, rather in any kind of income (Baumol, 1990), they try to escape the market mechanism – where profits are determined by competition –, if by doing so they can earn higher returns.

The attempt by entrepreneurs to escape the competitive pressure of the free market has been widely studied by the literature on public choice (Tullock, 1967), which has introduced the distinction between rent-seeking and profit-seeking behaviour (Buchanan, 1980). The rent-seeking activity, which manifests itself through corruption, lobbying, etc., chiefly aims to redirect policy proposal for the entrepreneur’s own advantage. In this way it is possible to earn permanent rents by altering the workings of the market process towards a non-competitive structure (Colander, 1984, p. 3), usually through the erection of legal barriers to entry (licences, quotas, tariffs, etc.). A feature peculiar to rent-seeking is that it has purely redistributive effects and the relevant literature has shown that, within a static context, this activity usually yields inefficiencies (Rowley, Tollison and Tullock, 1988).

The EGT has extended the study of rent-seeking to a dynamic context and has developed quite a few models which study the effects of rent-seeking on the growth process (Murphy, Shleifer and Vishny, 1991, 1993; Pecorino, 1992; Rama, 1993). This literature has reached the general but not unanimous, conclusion that, whatever the way in which rent-seeking is carried out, rent-seeking behaviour tends to depress the rate of growth of the economy; by contrast, profit seeking usually boosts growth.

It is our opinion that the idea that growth depends negatively on the existence of rent seeking can be traced back to classical economists, and in particular to Smith, although the latter maintains a more specific view concerning the nature of rent seeking compared with that of the EGT. We shall argue that Smith considers rent seeking as oriented only towards the creation of legal barriers to entry and that, according to this author, this activity, in this specific form, although not wasteful *in se*, negatively affects economic growth. Smith and the proponents of the EGT also hold different views as regards the mechanism through which rent-seeking affects growth, Smith's being much more elaborate that that of the EGT. Hence we can conclude that, from the explanatory point of view, the classical approach to growth – both in its original formulation due to Smith, Ricardo and Marx and in its modern rehabilitation due to the EGT – is a more satisfactory theory than the neoclassical one as it is able not only to explain the (sustained) growth of *potential* capacity but also to provide insights into the growth of *actual* capacity of economies.

The paper is organised as follows: the following section introduces the concept of rent-seeking behaviour. Section 11.3 provides a description of a simple model of endogenous growth in which entrepreneurs can pursue profit seeking or rent-seeking. Section 11.4 is devoted to surveying Smith's view of monopolistic markets and we shall point out that legal barriers to entry in Smith's analysis arise from 'rent-seeking' activity, and that Smith's view of the role of this activity on growth is very similar to the EGT view. Final conclusions are contained in Section 11.5.

11.2. COMPETITION AND RENT-SEEKING

The literature on rent-seeking starts with a paper by Gordon Tullock (1967), who questioned the calculation of the welfare losses of the static monopoly. He argued that the welfare losses are greater than the Harberger triangle since they should also include the monopolist's expenditure for gaining and maintaining the monopolistic position. Since then, the term rent-seeking has been variously defined (Brooks and Heijdra, 1989) with one of the most commonly used meanings being the following: rent seeking is 'the pursuit of profits via the use of government coercion' (Anderson, Rowly and Tollison, 1988, p. 100). By 'government coercion' is meant any activity by public authorities which alters the income distribution determined by the competitive market mechanism through redistributive actions (for example, taxing and government subsidies; Sturzenegger and Tommasi, 1994) or by creating legal barriers to entry like import licences or tariffs (Krueger, 1974; Bhagwati and Srinivasan, 1980; McCormick, Shughart and Tollison, 1984). In any case government intervention upon economic activity gives rise to rents benefiting particular groups. Therefore, if rent seeking yields higher returns than the usual productive activities, people compete through it by diverting resources from the usual productive activities. Thus, rent seeking is wasteful and creates inefficiencies.

Rent seeking may therefore be seen as just an alternative form of profit seeking: while profit seeking is carried out by agents seeking to maximise returns on their resources within the market system, rent seeking is carried out by agents who want to maximise the returns on their resources by exploiting the political power of coercion. The common opinion (Buchanan, 1980) is that profit seeking operates through innovation and the shift of resources towards employment with the highest returns. This yields, albeit unintentionally, positive effects on the economy because it ensures an efficient allocation of resources and generation of technical progress. Since profit-seeking activities are oriented to generating competitive advantage, they may create barriers to entry as well; however, since the market process does not ensure any protection, any barrier to entry usually turns out to be temporary and fades away through the competitive process of imitation. The political system, by contrast, is able to alter the distribution of the market mechanism by redistributive action or by creation of permanent legal barriers to entry that favour certain agents by allowing them to obtain, even in the long run, higher returns for their resources than their market levels. Rent seeking does not necessarily have beneficial effects on the economy, since it ensures neither technical progress nor efficient allocation of resources.

An alternative definition of rent seeking requires the employment of resources: rent seeking is 'any redistributive activity that takes up resources'

(Murphy, Shleifer and Vishny, 1993, p. 409). Private and public rent seeking is distinguished: ‘Private rent-seeking takes the form of theft, piracy, litigation, and other forms of transfer between private parties. Public rent seeking is either the redistribution from the private sector to the state, such as taxation, or alternatively from the private sector to the government bureaucrats who affect the fortunes of the private sector’ (Murphy, Shleifer and Vishny, 1993, p. 412). Private rent seeking has been widely used in recent years to explain the existence of the state, private property and markets (Grossman and Kim, 1995). We do not consider private rent seeking here since it is not helpful in explaining the role of rent seeking, as an alternative activity to market-based competition, in the process of growth. Hence, below we shall focus only upon public rent seeking.

11.3. RENT SEEKING AND GROWTH

Rent seeking, as analysed in the static framework and briefly surveyed in the preceding section, has been extended to a dynamic framework by the EGT. Particular attention has been focused on the effects on growth of the different ways in which rent seeking is carried out, for example by means of corruption (as in Barreto, 2000), or by means of lobbying (as in Mohtadi and Roe, 1998). In this section we shall provide a simple growth model with rent seeking. The model is deliberately as simple as possible in order to highlight the features of rent seeking and its effects on growth. Our analysis is an extension of the two-sector model of Murphy, Shleifer and Vishny (1991, 1993) to an explicitly growth context *à la* Harrod–Domar, as their model is essentially static.

Let us assume a two-sector economy with a continuum of agents indicated by the interval $I = [1, 2]$.[1] The two goods are indicated by 1 and 2 and either good can be used as consumption and as a capital good. Good 1 is the numeraire and, for the sake of simplicity, we assume that prices are fixed and that the price of good 2 is also equal to 1. Agent $a \in I$ is endowed with a positive amount of capital $K(a) = \overline{K}$, with the technology $Y_1(a) = \beta aK$ to produce good 1 and with the technology $Y_2(a) = \rho aK$ to produce good 2, where $\beta > 0$, $\rho \geq 0$. We also assume that $\beta > \rho$, which means that production of good 2 has a value-added lower than production of good 1. We also assume that production of good 2 is carried out together with a rent seeking activity which requires a fixed amount L ($0 < L < \overline{K}$) of capital. The way in which rent seeking affects the distribution of income is usually described in reduced form and is interpreted as due to a direct redistributive policy through, for example, taxing or subsiding (Murphy, Shleifer and Vishny, 1991,1993) or due to the creation of legal barriers to entry through, for

example, licenses or tariff barriers (Pecorino, 1992; Rama, 1993) in favour of sector 2. Therefore, rent seeking in sector 2 allows agent a to earn a rent equal to $R(a) = r(a)K$, where

$$r(a)=\frac{\delta a}{\int_{\Omega}\delta a d\mu}R(\Omega),$$

and where μ is the Lebesgue measure on I, Ω is the set of agents acting as rent-seekers, δ is a positive number and $R(\Omega)$ indicates the total rent extracted in the economy. Rent $R(\Omega)$ is assumed to be fraction t of the high value-added production of good 1, $0 \leq t \leq 1$. Thus, by employing capital K in sector 2, agent a yields a total income equal to

$$Y_2(a)+R(a)=\left[\rho+\frac{\delta}{\int_{\Omega}\delta a d\mu}R(\Omega)\right]a\left(\overline{K}-L\right).$$

Finally, we assume that each agent saves the fraction s ($0 \leq s \leq 1$) of his/her income and that capital does not depreciate.

Using the assumptions adopted, agents with a higher index are also the most efficient in both production of goods 1 and 2. Hence set Ω is always a left interval of point 2. Suppose, therefore, that agents in interval $[A, 2]$ produce good 2, hence total production in sector 1 is

$$Y_{T1}(A) = \int_1^A Y_1(a)da = \int_1^A \beta a\cdot\overline{K}da = \frac{(A^2-1)\beta}{2}\overline{K},$$

total production in sector 2 is:

$$Y_{T2}(A) = \int_A^2 Y_2(a)da = \int_A^2 \beta a(\overline{K}-L)da = \frac{(4-A^2)\beta}{2}(\overline{K}-L),$$

while total rent is

$$R_T([A,2]) = tY_{T1}(A) = t\frac{(A^2-1)\beta}{2}\overline{K}.$$

By using a^* to indicate the agent for whom the production of good 1 yields the same returns as the production of good 2, then a^* must satisfy the following equality:

$$(1-t)\beta a^*\bar{K} = \left[\rho + \frac{\delta}{\int_{[a^*,2]}\delta a d\mu} R([a^*,2])\right] a^*(\bar{K}-L);$$

i.e. $$(1-t)\beta a^*\bar{K} = \left[\rho + \frac{t\beta\bar{K}(a^{*2}-1)}{4-a^{*2}}\right] a^*(\bar{K}-L).$$

If $(1-t)\beta\bar{K} > \rho(\bar{K}-L)$ the above equation has only one solution in the interval [1,2]; therefore, every agent with an index higher than a^* will produce good 2, while every agent with an index lower than a^* will produce good 1.

An easy exercise of comparative statics shows that a^* decreases as t or ρ increases and L or β decreases. Therefore, the more effective the barriers to entry (measured by t and by L), the higher is the number of agents who turn to the production of good 2. In any case, the more talented people are those who turn first to rent-seeking whenever it yields higher returns than producing good 1. Therefore, rent seeking diverts entrepreneurial resources from the production of good 1 to the production of good 2. This has several implications at the macroeconomic level. In fact, the total production is:

$$Y_T(a^*) = Y_{T1}(a^*) + Y_{T2}(a^*) = \frac{(\beta-\rho)a^{*2}-(\beta-4\rho)}{2}\bar{K} - \rho L\left[\frac{4-a^{*2}}{2}\right].$$

Hence total production and savings, $sY_T(a^*)$, decrease as t increases and L decreases as we assume that $\beta > \rho$. Moreover, the total stock of capital employed in the production activity, $(a^*-1)\ \bar{K} + (2-a^*)(\bar{K} - L)$, decreases as a^* decreases. It follows, that the rate of capital growth,

$$sY_T(a^*)/\bar{K} = s\ \frac{(\beta-\rho)a^{*2}-(\beta-4\rho)}{2} - \rho\frac{L}{\bar{K}}\left[\frac{4-a^{*2}}{2}\right]$$

decreases as a^* decreases.

This model identifies a major general effect of rent-seeking on growth: it can make the high value-added activity unattractive by increasing the opportunity cost in carrying out that activity. Thus, it diverts capital from highly productive activities (production of good 1) towards lower productive activities (production of good 2) or totally unproductive activities (production of good 2 if $\rho = 0$). Moreover, the way in which the economy is modelled indicates that the human resources which turn to rent-seeking are usually the best resources available in the economy.

11.4. SMITH’S VIEW ON RENT SEEKING, COMPETITION AND GROWTH

In this section we show that according to Smith, rent seeking activities, competition and growth are closely related, as free competition maximizes the amount of resources devoted to the growth of wealth, while imperfect competition negatively affects the growth of wealth by reducing the amount of resources employed in productive investment, and the average rate of saving. While Smith has a narrower conception of rent seeking than the EGT, as he deals only with rent seeking oriented towards the creation of legal barriers to entry, he has a more complex view of the way in which rent seeking affects growth. However, Smith and the EGT share the view that rent seeking negatively affects growth.

According to Smith, the growth of a nation depends upon technology (‘division of labour’, WN I.i) and upon the stock of capital available which, in turn, depends upon the way in which labour is allocated between ‘productive’ and ‘unproductive’ employments (WN II.iii). Here we need just to recall that productive labour is any kind of labour which is employed to produce material goods, while unproductive labour is employed to produce services which, although useful, are not considered to contribute to the wealth of the economy (for details see Bowley, 1975, pp. 369–72, or Eltis, 1984, p. 78) and are substantially consumption activities (WN II.iii.6–7). In fact, while economic activities in which productive labour is employed yield profits (WN II.iii.6), those which employ unproductive labour do not yield profits and are supported just by income that has been produced elsewhere (WN II.iii.7). Unsatisfactory as it may be, the distinction between productive and unproductive labour is useful in that it highlights the Smithian idea that growth is determined by the way in which resources are employed between activities which increase income and activities which merely transfer it.

In chapter III of book II of *The Wealth of Nations*, after presenting the distinction between productive and unproductive labour as factors determining the accumulation of capital, Smith points out that the allocation of labour in productive and unproductive employments depends upon ‘parsimony’ and ‘frugality’ of individuals (WN II.iii.14); in fact:

> Parsimony, by increasing the fund which is destined to the maintenance of productive hands, tends to increase the number of those hands whose labour adds to the value of the subject upon which it is bestowed. It tends therefore to increase the exchangeable value of the annual produce of the land and labour of the country (WN II.iii.17).

By contrast,

> The prodigal… [by] diminishing the funds destined for the employment of productive labour, …, necessarily diminishes, …, the quantity of that labour which adds a value to the subject upon which it is bestowed, and, consequently, the value of the annual produce of the land and labour of the whole country (WN II.iii.20).

Prodigality in turn can be private and public. Actually, according to Smith, all public revenue is often employed in maintaining unproductive labourers (WN II.iii), although Smith recognises that this kind of employment may be useful (WN V.i). Moreover, mistakes in business can be assimilated to prodigality as both destroy resources which can be employed in productive uses (WN II.iii.26), although, in general, successful business compensates its unsuccessful counterpart (WN II.iii.27).

It is our opinion that, according to Smith, the competitiveness of markets, in particular the existence of legal barriers to entry, has a crucial role in determining the proportion between productive and unproductive labour, and therefore the growth of the economy. In order to support this view we now turn to the concept of competition developed in *The Wealth of Nations*.

The society that Smith has in mind in constructing his theory of value and distribution is a society divided into classes of resource owners: workers, capitalists and landowners, and the self-interest of each individual belonging to any of these classes is the origin of the competitive process. He conceives competition as a process carried on *within* the market by agents who desire to attain a particular goal. Smith uses this term mainly with two meanings: first, to describe the race amongst producers or consumers to sell or buy, respectively, the quantity of goods they have produced or they desire to buy; second, to describe the process of entry of resources into those markets in which they can secure higher remuneration. The former kind of competition, called price competition, ensures the attainment of the market price (WN I.vii.7–10). The latter kind of process, called intersectoral competition, is based upon the idea that individuals are greedy and seek to earn the highest remuneration for the resources they own through the market process.

The working of intersectoral competition can be hampered by possible impediments to the free movement of resources, i.e. by ‘barriers to entry’. In the presence of barriers to entry, market prices no longer gravitate around their natural values and the income of the resource owners who employ their resources in these sectors can be higher than the income of the same resource employed in competitive sectors even for long periods. Smith conceives of four causes of barriers to entry: (*i*) the scarcity of a particular resource, (*ii*) the secrets in trade (*iii*) the secrets in manufacturing and (*iv*) the privileges granted by law to an individual or to a trading company (WN I.vii.21–26). The first kind of barrier to entry will be called ‘natural’ and is due to the non-reproducibility of scarce natural resources. The second kind of barrier to

entry will be called ‘informational’ and is due to the imperfect information of the condition of demand, which allows producers to sell their product at a price higher than its natural level. The third kind of barrier to entry will be called ‘technological’ and is due to the existence of technical improvements which are kept secret and which allow the inventor to sell the goods at a higher market price, calculated with respect to the old technology, than the new natural price. The fourth kind of barrier to entry, which will be called ‘legal’, is due to the intervention of political authorities that, by law, are able to impede free movement of resources (for example, by means of monopoly, statute of apprenticeship, etc.).

Intersectoral competition has a crucial role in the classical theory of value and distribution as it ensures that, in competitive sectors, i.e. in those sectors in which there are no barriers to entry, market prices gravitate around their natural values (WN I.vii.15). Particularly important for our concern, however, is the role that free competition has on the allocation of resources: in fact, a main point of Smith’s view of the capitalistic process is that whenever the economic activity is left to the interest of individuals, the allocation of free mobile resources among different activity, governed by the ‘invisible hand’, turns out to be in accord with the ‘public interest’ (WN IV.ii.9). Smith never defines explicitly what the ‘public interest’ is. However it is reasonable to interpret it as the maximization of the wealth of nations. In fact, in the same chapter that contains the preceding statement Smith makes it clear that people, in seeking their own advantage and in competitive conditions, employ their resources ‘as near home as possible’ and in such a way that ‘its produce may be of the greatest possible value’ (Ibidem). By doing so, ‘every individual necessarily labours to render the annual revenue of the society as great as he can’ (Ibidem).

However, according to Smith not all barriers to entry are against the ‘public interest’. Indeed, while Smith deeply criticises legal barriers to entry, he seems to maintain a positive view of technological barriers to entry.[2] As for the latter, Smith gives an assessment of technological barriers to entry incidentally in Book V in discussing the only case in which he agrees to give a temporary legal monopoly to companies of merchants to establish new trade with high risk countries. In this case, Smith justifies the granting of a monopoly ‘upon the same principles upon which a like monopoly of a new machine is granted to its inventor, and that of a new book to its author’ (WN V.i.e.30). The justification for granting a temporary monopoly is that in all these cases the monopoly recompenses traders and researchers for the risks undertaken. However, Smith is very careful to emphasise that the monopoly must be temporary in order to ensure such beneficial effects on the public (Ibidem).

By contrast, as anticipated above, Smith expresses a radically different opinion with respect to permanent legal barriers to entry. In analysing trade with the colonies, Smith maintains that permanent monopolies affect negatively the growth of the economy as they reduce the amount of capital invested in productive labour and employ it in unproductive labour (WN IV.vii.c.49). Smith holds that legal barriers to entry, like technological barriers to entry, are created under the pressure of private interest. This is made clear in chapter X, where he points out that the resulting monopolies and corporations are established not only because of the sovereign's interest in raising money (WN I.x.c.17), but also under the pressure of mercantile interests:

> The government of towns corporate was altogether in the hands of traders and artificers, and it was the manifest interest of every particular class of them, to prevent the market from being over-stocked (WN I.x.c.18; see also IV.ii.21).

Legal barriers to entry, however, are not required only for protecting the incumbent firms from domestic potential competition:

> A great empire has been established for the sole purpose of raising up a nation of customers who should be obliged to buy from the shops of our different producers, … for the sake of that little enhancement of price which this monopoly might afford our producers, the home-consumers have been burdened with the whole expence of maintaining and defending that empire (WN IV.viii.53–4).

Again, this is the outcome of the lobbying of the capitalists on political institutions:

> It cannot be very difficult to determine who has been the contrivers of this whole merchantile system, not the consumers,…, but the producers, whose interest has been so carefully attended to, and among this latter class have been by far the principal architects (Ibidem. For other examples, see also WN IV.i.10, IV.iv.1 and IV.v.1).

From the above textual evidence it seems reasonable to conclude that Smith conceives the creation of legal barriers to entry as the outcome of a rent seeking activity deliberately carried out by landlords and entrepreneurs in their attempt to maximize rents or profits.[3, 4] However, some important differences must be emphasised between Smith's view and the modern view of rent-seeking. First, Smith conceives rent seeking as only directed towards the creation of barriers to entry. Moreover, consistent with part of the current literature on rent-seeking (see Anderson, Rowly and Tollison's definition of rent-seeking in Section 11.2), and in contrast with the alternative definition

(Murphy, Shleifer and Vishny's), it seems that Smith never explicitly considers the creation of legal barriers to entry as the outcome of a specific economic activity which employs resources. Rather, this activity is always conceived within a productive activity; hence, any activity aimed at the creation of legal barriers to entry is not wasteful *in se*. However, like the EGT, Smith believes that legal barriers to entry do have *negative* effects on income and growth, although, as we shall see immediately, he has a more complex view than the EGT as far as this relationship is concerned.

In chapter VII of book IV of *The Wealth of Nations* Smith provides a detailed account of the effects of legal barriers to entry – in particular the monopoly granted on colonial trade – on the growth of an economy. Apart from specific negative effects that monopoly on trade produces due to the peculiar nature of the barrier to entry, he pinpoints two general negative effects, one direct and the other indirect, that monopolies have on the stock of capital employing productive labour. The direct negative effect is due to the fact that monopolies, by impeding entry of new capital in markets, contract the stock of capital employing productive labour; the indirect negative effect is due to the fact that monopolies reduce prodigality. The first effect supports Smith's general view that barriers to entry divert resources from their 'natural' allocation, which is the competitive one and which ensures a higher level of productive labour with respect to the monopolistic case (WN IV.ii). Thus, the presence of a monopoly reduces the amount of capital in the country that employs productive labour. This implies that the income generated in the country decreases, with a consequent decrease in the savings and in the rate of capital accumulation (WN IV.vii.c.57). Moreover, Smith is very careful to stress that this contraction affects equally all kinds of income, with the exception of the profits in monopolistic sectors. Indeed, a reduction in the productive labour employed implies a reduction in wages (Ibidem). Furthermore, the increase in the rate of profits in monopolistic sectors, by depressing the remuneration of capital in other sectors relative to the monopolised sectors, discourages improvement in land and this, in turn, 'necessarily retards the natural increase of another great original source of revenue, the rent of land' (WN IV.vii.c.58. Finally, although monopoly yields an increase of profits in the monopolised sector, Smith believes that total profits are lower than in the competitive situation because of the entry constraints of capital in monopolised markets (WN IV.vii.c.59).

As already stated, the indirect effect of monopolies on growth is due to the fact that they negatively affect parsimony (WN IV.vii.c.61). However, the reduction of savings is not caused only by the misbehaviour of capitalists, but by a generalised reduction in the saving ratio as the monopolist's behaviour will be imitated by all the other classes (Ibidem). Finally, Smith explicitly recognises rent-seeking activities directed towards the creation of legal barriers to entry. Moreover, he also holds the idea that rent-seeking activity is detrimental to the growth of an economy, directly through the reduction in the stock of capital in the economy and the consequent reduction of productive labour employed and of the saving rate, and indirectly through destruction of parsimony. Moreover, he explicitly recognises that the increase in profits in monopolistic sectors may reduce the incentive of investing in activities that improve technology.

The direct effect of legal barriers to entry on the stock of capital can be illustrated graphically as follows.[5] Consider an economy with two sectors, say agriculture and industry. In Figure 11.1 the segment *O–O'* measures the total stock of capital available in the economy. Curve Π_i indicates the rate of profit in industry as a function of the stock of capital employed in that sector, while curve Π_a, which should be read from origin *O'*, indicates the rate of profit in agriculture as a function of the stock of capital employed in that sector. These curves are decreasing because Smith believes that the market price, and therefore, the rate of profit decrease as the production level increases. Intersectoral competition ensures that fraction $\overline{OA}$ of capital should be employed in industry and that fraction $\overline{O'A}$ should be employed in agriculture so that the common rate of profit $r^* = r_a = r_i$ is generated. Following Smith, r^* is also the rate of interest (WN I.ix.4) so no production can yield in the long run a rate of profit lower than r^*. If a legal barrier to entry is introduced in industry which ensures that the producers earn r_i', then a contraction of amount $\overline{AB}$ of stock in industry must occur. Eventually, the total stock of capital in the economy will be equal to $\overline{OB}+\overline{O'A}$, where $\overline{OB}$ is employed in industry and $\overline{O'A}$ in agriculture (the competitive sector).

As seen above, Smith also suggests that if agriculture yields lower returns than those earned in the monopolistic sector, then capital in that sector may be withdrawn, thus causing reduction of rents. If this occurs, capital will also be reduced in agriculture, and possibly to such a level that ensures a new, higher, common rate of profit $r' = r_i' = r_a'$. If we interpret curve Π_a as illustrating the rate of profit on the stock of capital employed also for improvements in agriculture,[6] the amount of capital that must be withdrawn in agriculture in order to ensure the common rate of profit r' is equal to $\overline{AB'}$. Now the total amount of capital employed, therefore, is $\overline{OB}+\overline{O'B'}$.[7]

Smith's view of the effects of rent seeking on growth via prodigality can be easily formalised by means of the model introduced in the preceding section. The Smithian growth model with legal barriers to entry due to rent seeking makes the following assumptions: (*i*) $L = 0$; (*ii*) $\beta \geq \rho$; and (*iii*) the rate of saving of producers in sector 1, s_1, is greater than the rate of saving of producers in sector 2, s_2. Assumptions (*i*) and (*ii*) should capture the idea that for Smith rent seeking is not necessarily wasteful and it can be associated with productive activities. Assumption (*iii*) should reflect Smith's idea

concerning the prodigality of rent seekers. It can be noted that in Smith's model the rate of growth of capital is:

$$\frac{sY_T(a^*)}{\overline{K}} = \frac{(s_1 - s_2)a^{*2} - (s_1 - 4s_2)}{2}\beta,$$

thus, because of assumption (*iii*), it decreases as the production of good 2 yields higher returns (a^* decreases).[8]

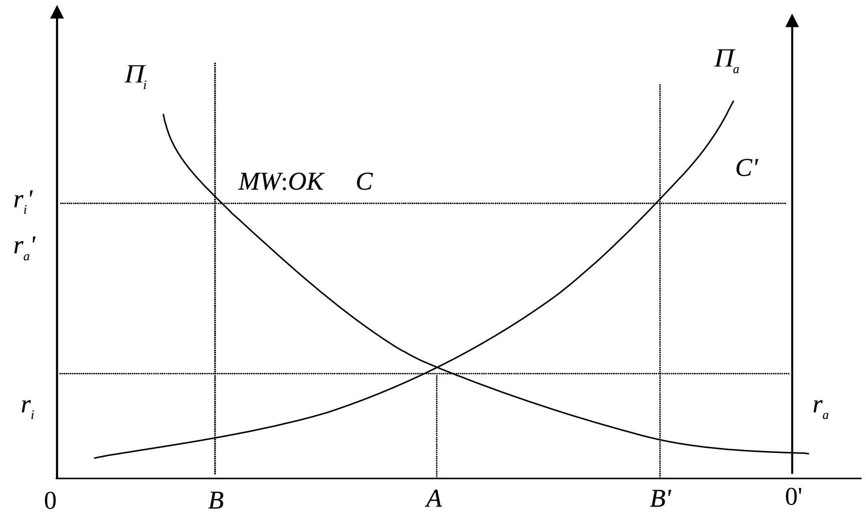


Figure 11.1

11.5. CONCLUSIONS

We have shown that the EGT and Smith share the view that legal barriers to entry created through rent seeking negatively affect growth because they tend to increase the opportunity cost of pursuing productive activities, and therefore tend to employ resources in 'unproductive' or 'less' productive activities. We have also pointed out that there are differences between the EGT and Smith's views concerning the mechanism through which rent seeking affects growth: unlike the EGT, according to Smith rent-seeking is directed towards the creation of legal monopolies, and moreover it is not *in se* a wasteful activity as it is usually carried out within a productive activity. Finally, it is through the creation of such barriers to entry that, according to Smith, rent seeking has negative effects on growth. He identifies two negative effects on growth; the first, direct, reduces the amount of the overall (productive) capital employed in the economy because barriers block entry and because, as stated above, it increases the opportunity cost in pursuing productive activities; moreover, it reduces the aggregate magnitude of each kind of income. The second, indirect effect entails the reduction in parsimony.

NOTES

* I would like to thank the participants at the workshop organised by the MURST group ‘Classical Themes in Modern Growth Theory’ held in Rome 1–4 March 2001 for their useful comments. I am indebted to Mario Lavezzi and Neri Salvadori for insightful suggestions on a draft of this paper. The financial support of the MURST (Cofin 1999) and the University of Catania (Fondi di Ateneo) is also gratefully acknowledged.

1. Murphy, Shleifer and Vishny’s 1991 model is a special case of this model by setting $\rho = 0$ and L = $\bar{K}$, i.e. they conceive rent-seeking as a purely redistributive activity.
2. Smith seems to have expressed no specific opinion as far as natural and informational barriers to entry are concerned.
3. For an interpretation along this line and for a more detailed analysis of the relationship between private interest and public intervention see Stigler (1975).
4. Interestingly, Smith also considers that political authorities are petitioned not only to create legal barriers to entry, but also to increase rivals’ costs (WN I.ix.5).
5. The approach we adopt is essentially a partial equilibrium one. Although it is not consistent with Smith’s overall analysis, which is of a general equilibrium nature, it seems to be what he has in mind when he concludes that ‘[t]he increase of stock, …, tends to lower profits’ (WN I.ix.2: see also II.iv.8).
6. In this case the reduction in the rate of profit as the capital stock increases is due not only to the consequent reduction in the price of produce due to increased production, but also to the reduction of the returns of improvements.
7. Smith also holds that the total amount of profits decreases, even if the rate of profit increases. This means that according to Smith in Figure 11.1 area $Or_i r_a O'$ is larger than area $Or_i'CB + B'C''r'_a O'$.
8. A formal analysis of the effects of rent seeking on growth through reduction of the total stock of capital in productive activities seems to be a difficult task since it should be carried out within a general equilibrium analysis rather than a partial equilibrium one like that employed by Smith (see footnote 5).

12. R&D models of economic growth and the long-term evolution of productivity and innovation

Mauro Caminati

12.1. INTRODUCTION

The ratio between the number of scientists and engineers engaged in research and development (R&D) and the level of total employment increased dramatically in the U.S.A. in the second half of the twentieth century. Let us call h_L the ratio between employment outside of R&D and total employment. In the U.S.A. $(1 - h_L)$ was nearly three times as large in 1993 as in 1950, with a pronounced upward fluctuation in the period 1960–70 due to government-funded R&D. Jones (2002) estimates that from 1950 to 1993 there was an even larger rise in the researchers/employment ratio in the set of G-5 countries (France, West Germany, Japan, the United Kingdom and the United States).

It is quite striking that the observed dramatic rise of R&D employment did not show up in the productivity figures. As is well known, the growth rate of GDP per hour tended to decline in the advanced countries after the 1950–70 ‘golden age’. The decline was less pronounced in the U.S.A. because this country did not enjoy the boom in productivity from technological catching up after the Second World War. Hence the U.S. experience provides a more telling indication of the relation between R&D effort and productivity growth for a country located on the frontier of technological knowledge.[1]

I will refer the mentioned rise in the researchers/employment ratio as stylised fact (*a*) and the relatively constant growth rate of GDP per hour in the second half of the twentieth century as stylised fact (*b*).[2]

The question discussed in this paper is how the R&D models developed within the recent revival of general-equilibrium-growth theory cope with the facts (*a*) and (*b*).[3] A similar question was addressed in an influential paper written by C.I. Jones and published in 1995. Jones observed how the R&D growth models developed to that date displayed a ‘scale effect’ of the number of researchers on the growth rate of GDP per capita. These models are criticised by Jones because the 'scale effect' is in striking contrast with the evidence. In the same paper he builds a model, which he defines *semi-endogenous*, where innovations are still the outcome of purposeful and costly R&D effort, but the steady-state growth rate of output per capita is completely determined by the technological parameters and the rate of population growth. It is therefore independent of the level of population, of preferences, and of policy variables that do not affect technology. The family of R&D growth models with these properties is called here *non-endogenous*. By contrast, the endogenous R&D models of general-equilibrium growth are those where per-capita GDP growth depends upon preferences and/or policy variables generally. The basic structure of the endogenous and non-endogenous general-equilibrium models of economic growth is discussed in Sections 12.2, 12.3 and 12.4.

Partly as a reaction to Jones' critique, a second generation of endogenous R&D growth models appeared in the late 1990s. In this second generation, besides 'intensive' innovations that increase the productivity of the intermediate good produced in their sector of application, there are 'extensive' innovations, that increase the number of intermediate goods. In steady-state equilibrium, the number of intermediate goods (hence of sectors) grows at the population growth rate n, so that, in steady state, the number of intensive-researchers per sector is constant. This implies that the 'scale effect' on the rate of growth disappears. In other words, there is a dilution of the 'scale effect' across the growing number of intermediate-good sectors.

The steady-state predictions of the second-generation endogenous and also of the non-endogenous R&D growth models are still at odds with the evidence presented at the beginning of this introduction. The dramatic long-term rise of the R&D employment share $(1 - h_L)$ reveals that the long-term growth path of the U.S. economy cannot find a theoretical approximation through the hypothesis that the economy has been growing in the neighbourhood of a single steady-state path.[4] The observed long-term rise of $(1 - h_L)$ and the approximately constant rate of productivity growth may be more consistent with the hypothesis of a transitions path induced by exogenous parameter changes. This issue is addressed in Section 12.5.1. My conclusion here is that the changes in the technological parameters required to reconcile the stylised facts (*a*) and (*b*) above may be implausibly large.

In Section 12.5.2 I suggest that the failure of the R&D growth models to reconcile the constant productivity growth with the long-term rise of the R&D employment share can be interpreted as the result of technological assumptions that make abstraction from complementarity in production. Following a system-like view of technology which owes much to the contributions of Nathan Rosenberg, Joel Mokyr and many others and which

can be traced back to Karl Marx, I stress the relation between the arrival of new technologies and the growth of variety and show how this may help to explain the stylised facts (*a*) and (*b*).

The focus of this paper is on steady-state results. When transitions between different steady states are involved, e.g. in Section 12.5.1, the implicit assumption is that transitional dynamics are monotonic. Eicher and Turnovsky (1999b and 2001) show that non-monotonic transitional paths may exist in the non-endogenous growth models with two endogenously accumulating factors, knowledge A and capital K. In what follows the endogenously accumulating factors are capital K, intensive technical knowledge A and extensive technical knowledge N. A general analysis of the transition dynamics for the R&D growth models of this type is still lacking.[5] The discussion of how it may be relevant to the theme of this paper is left to future work.

An important caveat must be added. In what follows, the rigid supply orientation of the general-equilibrium models of economic growth is taken for granted and is not questioned. This is not because the author is not aware of the biases that are introduced when co-ordination problems or stability in the disequilibrium dynamics are disregarded. These issues are simply outside the scope of this paper.[6] In a similar vein, the paper is unconcerned with the criticism that may be levelled at the use of capital aggregates in theoretical models. The attitude is simply to take the model predictions for what they are and discuss their consistency with broadly defined stylised facts.

12.2. A UNIFYING REPRESENTATION OF TECHNOLOGY

In what follows I build a framework which embeds different views of the relation between output growth and the generation of new inputs, as may be encountered in R&D growth models. This is done under a number of simplifying assumptions about technology that still enable us to discuss usually neglected issues, such as the role of complementarities and the relation between technological compatibility and knowledge spillovers. The main simplifying assumption is that the service characteristics of final output Y are unchanged throughout, that Y can be either consumed or accumulated in the form of capital and that it is produced by means of intermediate goods and labour. The number of available intermediate goods N_t changes through time as a result of innovation activities.

Assume that the number of service-characteristic types that exist in nature is finite. An intermediate good is a pair $(v, A_v) \in R_+^2$, where v is the intermediate–good variety (which identifies a class of functions performed by the good, that is, a composition of the associated flow of service characteristics) and A_v is the technological level, or generation, to which (v, A_v) belongs. In principle, we should expect A_v to have only an ordinal meaning, possibly with the further ordinal implication that later generations of a variety are also more productive. This is not, however, the interpretation we find in the new-growth literature, where A_v is an index leading to a cardinal productivity measure. The marginal product of (v, A_v) is a known time-invariant function of A_v (and possibly other variables). This leads to a time-invariant production possibility frontier, describing the productive potential of every possible present and future combination of intermediate goods.

12.2.1. Production of Material Goods

Final output Y is produced by means of intermediate goods and labour by perfectly competitive firms which, individually, face constant returns to scale. Following the R&D growth literature, we introduce a set of simplifying assumptions implying that at every date t only the highest (and latest) available technology level $A_{v,t}$ of each variety v is used. This will be the case to the extent that the value of the productivity gain from using the latest generation of a given variety invariably dominates the cost differential associated with the same choice. [7] In fact, these models assume a particular substitutability relation between intermediate goods, to the effect that they enter the production function in an additively separable form. Recalling our simplifying assumptions, the individual production function is:

$$Y_t = N_t^{\gamma} L_{Y,t}^{1-\alpha} \left[\int_{v=0}^{N_t} A_{v,t}\; x_{v,t}^{\alpha}\, \partial v \right] \tag{1}$$

where x_v is a quantity of the intermediate-good variety v and L_Y is labour employment in the production of final output. It follows from (1) that the marginal product at t of the intermediate good $(v, A_{v,t})$ is $\alpha\, N_t^{\gamma} L_{Y,t}^{1-\alpha}\; A_{v,t}\; x_{v,t}^{\alpha-1}$. It is independent of the inputs of the other intermediate goods, although it may depend, if $\gamma \neq 0$, on the total *number* of intermediate goods cooperating with it. The above form of independence is interpreted here as resulting from the lack of technological complementarity between any two intermediate goods.

Intermediate goods are produced by local monopolists through a different set of activities. The reason why firms in the intermediate-good sector cannot be perfectly competitive is quite robust (Arrow, 1987 and 1998; Romer, 1990). The right to produce a new intermediate good involves an innovation cost that represents a fixed cost, because once the knowledge to produce a

unit of a new good is acquired, it can be applied to the production of an indefinite number of units. If intermediate-goods production is otherwise subject to constant variable costs, we are faced with a clear case of increasing returns.

The input of the activity for producing one unit of (v, A_v) is a quantity of capital K which depends positively on the technology level A_v. K units of capital invested in the production of good (v, A_v) give rise to K/A_v^{ω} units of this good, where $\omega > 0$, thus implying that more capital intensive methods are required to produce intermediate goods of a later generation. Hence $K_v = x_v A_v^{\omega}$. Howitt (1999) adopts a similar increasing-capital-intensity assumption and claims that capital used in intermediate-goods production can be interpreted as human capital. The above specification implies that the average and marginal cost, in terms of final output, of producing (v, A_v) is rA_v^{ω}, where r is the rental price of capital. Since we abstract from depreciation, r is also the rate of interest.

The monopoly profit from producing $x_{v,t}$ is:

$$\pi_{v,t} = \alpha(1-\alpha)\; N_t^{\gamma} L_{Y,t}^{(1-\alpha)} A_{v,t} x_{v,t}^{\alpha}$$

Aghion and Howitt (1998, ch. 12) and Howitt (1999) obtain a monopoly output which is uniform across varieties and independent of A,[8] by setting $\omega = 1$. We hold to the latter simplifying assumption to obtain:

$$x_{v,t} = \alpha^{\frac{2}{1-\alpha}} N_t^{\frac{\gamma}{1-\alpha}} L_{Y,t} r_t^{\frac{1}{\alpha-1}} = x_t \tag{2}$$

In equilibrium, final output Y is then:

$$Y_t = N_t^{\gamma} L_{Y,t}^{1-\alpha} N_t A_t x_t^{\alpha} = \alpha^{\frac{2\alpha}{1-\alpha}} N_t^{\frac{1-\alpha+\gamma}{1-\alpha}} L_{Y,t} r_t^{\frac{1}{\alpha-1}} A_t \tag{3}$$

where A_t is the average technology level across intermediate goods:

$$A_t = \frac{1}{N_t} \left[\int_{v=0}^{N_t} A_{v,t}\, \partial v \right]$$

An equivalent equilibrium expression of Y_t is obtained by observing that, if h_K is the capital share employed in material, as opposed to knowledge, production, then in equilibrium we must have $(h_{K,t} K_t) / A_t = N_t x_t$. Hence:

$$Y_t = N_t^{\gamma} \left(h_{L,t} L_t \right)^{1-\alpha} N_t^{1-\alpha} A_t^{1-\alpha} (h_{K,t} K_t)^{\alpha} \tag{4}$$

It is then clear how the assumption $\gamma = \alpha - 1$ (see, for instance, Aghion and Howitt, 1998, ch. 12) sterilises the effects of the growing number of varieties on final output, which result from the additively separable way in which the single varieties enter the production function. Where these effects are not sterilised, because $(1 - \alpha + \gamma) > 0$, the production function corresponding to a constant technology level contains a form of increasing returns due to specialisation, as measured by N. The best known example along these lines is probably Romer (1990), which assumes $\gamma = 0$.

Recalling that in steady state the rate of interest is constant, and the labour and capital shares employed in the (final and intermediate) output sector are also constant, equation (2) yields the steady-growth equation:

$$g_Y = g_L + \frac{1-\alpha+\gamma}{1-\alpha} g_N + g_A$$

where g_i is the proportional instant rate of change of variable i. In particular, if following Romer (1990) we set the restrictions $\gamma = 0$ and $g_A = 0$, the above relation boils down to $g_y = g_L + g_N$, where it is apparent that the growth rate of per-capita output is simply the growth rate in the number of specialised varieties.

12.2.2. Intensive Innovations

An intensive innovation in sector v arriving in the interval $t + \mathrm{d}t$ is the stochastic outcome of the innovation effort performed at t in this sector. The innovation contributes to shifting the technology frontier according to

$$\dot{A}_{t\,Max} = \frac{\delta}{N_t} A_{t\,Max} \tag{5}$$

and brings $A_{v,\,t}$ to the shifted frontier. Thus, access to the frontier technology level is available, but not costless, to every successful *intensive* innovator operating in sector v. The knowledge increment has elasticity +1 with respect to $A_{t\,Max}$ and elasticity $-$ 1 with respect to the number of sectors in the economy (Aghion and Howitt, 1998, ch. 12). The idea is here that the higher the number of sectors, the lower the impact of an innovation in sector v on the technology frontier.

The Poisson arrival rate of an intensive innovation in sector v at t is:

$$\phi_{v,t} = \lambda (u_{L,v,t} L_t)^{\theta} (u_{K,v,t} K_t)^{\xi} A_{t\,Max}^{\chi} \tag{6}$$

where $\xi \geq 0$, $\theta > 0$, λ is a constant, $u_{L,v}$ and $u_{K,v}$ are the fractions of total labour and capital invested in intensive R&D on variety v.

The returns offered by the investment of *rival*-resources in intensive R&D are constant or decreasing, depending on $\theta + \xi = 1$ (Barro and Sala-i-Martin, 1995, ch. 7), or $\theta + \xi < 1$. The second case arises if there is a congestion effect on the returns to R&D investment (Stokey, 1995; Howitt, 1999), with the result that the larger the rival resources invested in research, the higher the probability that independent innovation efforts produce the same outcome.

The parameter χ is meant to capture how the arrival rate is affected by the frontier knowledge stock $A_{t\,Max}$. There are two main forces at work here, which act in opposite directions. Thus, we may split the parameter χ into two components, $\chi = \chi_1 + \chi_2$. χ_1 is the so called 'complexity effect': more advanced technology levels are progressively more difficult to discover as a result of the increasing complexity of the search activity. Thus, we have $\chi_1 < 0$. This is the assumption we find in a number of search models of R&D-based economic growth (Jovanovic and Rob, 1990; Stokey, 1995; Kortum, 1997).[9] The parameter $\chi_2 > 0$ captures the 'standing on giants' shoulders' effect (see Merton, 1965; see also Caballero and Jaffe, 1993), which postulates that a higher frontier knowledge increases the probability of invention because an investment in intensive R&D creates the opportunity to exploit a knowledge spillover from the technology frontier to the innovators. This positive influence of knowledge on the innovation-success probability is distinct from and indeed adds to the influence of the stock of ideas on the size of the knowledge shift, which takes place *if* the innovation arrives (see (5) above). To this extent, it is unclear what are the grounds for assuming that the giants' shoulders effect is positive and is close in absolute magnitude to the complexity effect. We shall see nevertheless that the restriction $\chi = \chi_1 + \chi_2 = 0$ (or other equivalent condition) is characteristic of the R&D endogenous-growth models.

The main simplifying hypothesis introduced with (6) is that the success probability of intensive R&D on variety v is independent of the distribution of the local stocks $A_{v,t}$. Together with (2) this implies that the intensive research effort and the arrival rate are uniform across sectors. Other formulations (see, for instance, Barro and Sala-i-Martin, 1995, ch. 7) relate the complexity effect and the giant's shoulders effect for sector v to the local stock $A_{v,t}$. The same property of a uniform equilibrium arrival rate is however imposed also in this case, by means of ad hoc restrictions introduced to this end.

Since each agent engages in R&D independently of the agents in the same or in other sectors and the equilibrium research effort is uniform across sectors, the aggregate rate of intensive innovations is deterministic and equals

$$N_t\phi_{v,t} = N_t\lambda(u_{L,v,t}L_t)^{\theta}(u_{K,v,t}K_t)^{\xi}A^{\chi}_{t\,Max} = N_t\lambda\left(u_L\frac{L_t}{N_t}\right)^{\theta}\left(u_{K,}\frac{K_t}{N_t}\right)^{\xi}A^{\chi}_{t\,Max} \quad (7)$$

where u_L and u_K are the aggregate labour and capital shares invested in intensive R&D.

12.2.3. Extensive Innovations

An 'extensive' innovation is the introduction of a new variety v. On the assumption that there is an external effect such that the technical knowledge in the economy affects the technology level of a new variety, a-not-too-implausible restriction is that the technology level distribution of a new variety corresponds to the technology level distribution across the existing varieties (Howitt, 1999). This implies that extensive innovations at t do not affect the average technology level in the economy A_t. An assumption to the same effect is that new varieties arriving at t have a deterministic technology level A_t (Peretto, 1998).

We assume that the extensive innovation effort is related to the creation of new varieties by the deterministic law:

$$\dot{N}_t = \beta(z_{L,t}L_t)^{\varepsilon}N_t^{\tau}(z_{K,t}K_t)^{\psi}A_t^{\nu} \equiv \phi_{N,t} \quad (8)$$

where β is a constant and z_L is the fraction of total labour employed in extensive R&D. We impose the restriction $\varepsilon > 0$, $\psi \geq 0$, $\tau \geq 0$, $\nu \geq 0$. The case $\varepsilon + \psi < 1$ indicates that there are decreasing returns with respect to the scale of the rival resources invested in extensive search. The restriction is referred to as the 'congestion hypothesis'. A positive τ bears the interpretation that a higher number of varieties amounts to a wider knowledge base in the economy as a whole and therefore facilitates the discovery of yet new varieties. If this is in itself quite plausible, far more questionable appear to be 'point restrictions' such as $\tau = 1$, or $\tau = 0$, as may be found, for instance, in the pure variety-extension model of Romer (1990) and in Peretto (1998), respectively.

The parameter ν indicates how the production of an extensive innovation flow $\dot{N}$ of technology level A is related to the size of the average technology index A. If $\nu = 0$, then the cost (in terms of rival resources invested in extensive R&D) of producing a given innovation flow $\dot{N}$ with average technology level A is independent of A (Peretto, 1998). If $\nu > 0$ (< 0), this cost is decreasing (increasing) in A. The restriction $\nu > 0$ fits with the idea that the growth of technical knowledge along the quality dimension goes hand in hand with a growing 'complexity' of technology, which has a positive effect on the ease with which new varieties are discovered.[10]

12.3. STEADY-GROWTH EQUATIONS

A steady state, or balanced-growth path, is a particular constant-growth path such that the growth rate of every variable is constant *for ever*. Since the employment shares of the factors cannot exit the interval [0, 1], the definition immediately implies that the growth rate of such variables is zero on a balanced path.

The assumptions of Section 12.2.2 imply that the ratio $(A_{t\,Max} / A_t)$ converges to $(1 + \delta)$ (see Aghion and Howitt, 1998, p. 412). Assuming that convergence has already taken place, from (5) and (7) we obtain the following shift in the average technology level at time t, resulting from the intensive R&D in the N sectors:

$$\dot{A}_{t,Max} = \delta\lambda\left(u_{L,t}\frac{L_t}{N_t}\right)^{\theta}\left(u_{K,t}\frac{K_t}{N_t}\right)^{\xi}A_t^{\chi+1} \quad (9)$$

Recalling that on a constant-growth path $\dot{A}_t$ and A_t grow at the same rate, using (4), (8) and (9) we write the steady-state growth equations:

$$g_A[-\chi] + (\xi + \theta)\,g_N - \xi\,g_K = \theta\,n \quad (10)$$

$$-\nu\,g_A + (1 - \tau)\,g_N - \psi\,g_K = \varepsilon\,n \quad (11)$$

$$-(1 - \alpha)\,g_A - (\gamma + 1 - \alpha)\,g_N + g_K(1 - \alpha) = (1 - \alpha)\,n \quad (12)$$

If we define the variables $k \equiv K/N$, $l \equiv L/N$, so that

$$g_K = g_k + g_N,\; n = g_l + g_N,$$

then (10), (11) and (12) yield the following system:

$$\begin{bmatrix} -\chi & 0 & -\xi \\ -\nu & 1-\tau-\varepsilon-\psi & -\psi \\ -(1-\alpha) & -(\gamma+1-a) & 1-a \end{bmatrix}\begin{bmatrix} g_A \\ g_N \\ g_K \end{bmatrix} = \begin{bmatrix} \phi g_1 \\ \varepsilon g_1 \\ (1-a)g_1 \end{bmatrix} \quad (13)$$

12.3.1. Endogenous R&D Growth

Let $[I - \Gamma]$ be the square matrix in the left-hand-side of (13). If $[I - \Gamma]$ has a non-zero determinant, the steady-state growth rates of A, N and K are fully determined by equations (13), hence by technology, given the exogenous growth rate of population. Thus Det $[I - \Gamma] \neq 0$ states that preferences do not have any bearing on the speed of steady-state growth and policy measures by a government are equally ineffective, unless they are able to affect the technological parameters. It is then apparent that the crucial assumption of the endogenous R&D growth models is Det $[I - \Gamma] = 0$. In this case the coefficients in (13) are linearly dependent and additional equations are necessary to determine the steady-state growth rates of the variables. One missing equation is derived from the first-order conditions associated to the utility-maximisation problem:

$$Max: \int_{t=0}^{\infty} \frac{c_t^{1-\sigma}-1}{1-\sigma} e^{-(\rho-n)t} \partial t$$

subject to the flow budget constraint that per-capita consumption at t c_t is not negative and is constrained by wage and interest income minus the accumulation of stocks at *t* (see Barro and Sala-i-Martin, 1995, ch. 2). ρ is the rate of time preference and $(1/\sigma)$ is the constant inter-temporal elasticity of substitution. In particular, the proportional growth rate of c_t must satisfy:

$$g_c(t) = \frac{r_t - \rho}{\sigma}$$

where *c* is per capita consumption. In steady state $n + g_c = g_Y = g_K$.

The restriction Det $[I - \Gamma] = 0$ may be of course introduced in a number of ways. The standard practice of endogenous growth models with intensive R&D is to postulate the special case: $\chi = 0$ and $\xi = 0$ (see, for instance, Grossman and Helpman, 1991; Aghion and Howitt, 1992; Howitt, 1999; Peretto, 1998; Young, 1998; Barro and Sala-i-Martin, 1995; ch. 7). This is the case considered in the sequel of Section 12.3.1, yielding:

$$\frac{\dot{A}_t}{A_t} = \delta\lambda \left(u_{L,t} \frac{L_t}{N_t} \right)^{\theta} \tag{14}$$

As is also revealed by the first equation of system (13), with $\chi = \xi = 0$ consistency with steady state requires $g_N = n$, that is, $g_l = 0$. In particular, in the models where extensive innovations are not contemplated, so that *N* is constant, it is assumed that *L* is also constant and there is a scale effect of the intensive-research employment level on the growth rates of *A* and *Y*. This occurs in the pure quality expansion model of Grossman and Helpman (1991), Aghion and Howitt (1992) and Barro and Sala-i-Martin (1995, ch. 7). Jones (1995) draws attention to the lack of empirical corroboration for the hypothesis of a scale effect on the growth rate. In models with a growing

population, equation (14) is reconciled with the lack of any scale effect on the steady-state rate of growth, by introducing special assumptions which make sure that L / N is constant (Howitt, 1999), or at least converges to a fixed steady-state value (Peretto, 1998; Young, 1998). With the simplified specification of equation (8) considered below ($\nu = 0$, $\psi = 0$), the required restriction is $\tau + \varepsilon = 1$. This implies:

$$\frac{\dot{N}_t}{N_t} = \beta z_{L,t}^{\varepsilon} \left(\frac{L_t}{N_t} \right)^{\varepsilon}$$

and using the steady-state condition $g_N = n$, this yields

$$mz_L = \left(\frac{n}{\beta} \right)^{\frac{1}{\varepsilon}} \tag{15}$$

where *m* is the steady-state value of *L*/*N*. There are two different sets of steady-state solutions of the endogenous model, as specified above, which correspond to the possibility that: (*i*) the costs of one additional unit of labour effort invested in extensive or intensive R&D are identical; (*ii*) these costs are not identical. Case (*i*) is considered in the next section, case (*ii*) in appendix A.

We shall proceed under the further simplifying assumption $\gamma = \alpha - 1$ (see equation 12), so that $g_K = g_A + n$. Thus $g_c = g_A = (r - \rho)/\sigma$.

12.3.1.1. Identical opportunity cost of effort in extensive and intensive R&D

Suppose the only cost of one additional unit of labour effort in extensive or intensive research is the forgone opportunity of obtaining the wage rate *w* by selling that unit in the labour market. Free entry in research implies that, if the equilibrium levels of intensive and extensive R&D are positive, then the private instantaneous marginal returns from innovation effort must be identical between the two activities and must be equal to the wage rate *w*.

$$\left[\frac{\phi_{v,t}}{u_{L,v,t} L_t} \right] v_{v,t} = \lambda \left(u_{L,t} \frac{L_t}{N_t} \right)^{\theta-1} V_t = w_t = \left[\frac{\phi_{N,t}}{z_{L,t} L_t} \right] V_{N,t} = \beta \left(z_{L,t} \frac{L_t}{N_t} \right)^{\varepsilon-1} V_{N,t} \tag{16}$$

where $V_{v,t} = V_t$ is the expected value of a quality innovation in any sector *v* at time *t*, and $V_{N,t}$ is the expected value of an extensive innovation at time *t*. Moreover, with our production function (4) we have:

$$w = (1 - \alpha) h_L^{-\alpha} q^{\alpha} A \tag{17}$$

where $q \equiv K/AL$.

Let $v_t \equiv V_t / A_{t,Max}$ and $v_{N,t} \equiv V_{N,t}/A_t$; in other words, v_t and $v_{N,t}$ are the productivity adjusted values at time t of an intensive and extensive innovation, respectively. From (16) and (17):

$$v_t = \frac{1-\alpha}{\lambda(1+\delta)} h_L^{-\alpha} q^{\alpha} \left(u_{L,t} \frac{L_t}{N_t} \right)^{1-\theta} \tag{18}$$

$$v_{N,t} = \frac{1-\alpha}{\beta} h_L^{-\alpha} q^{\alpha} \left(z_{L,t} \frac{L_t}{N_t} \right)^{1-\varepsilon} \tag{19}$$

Moreover, one obtains the asset equations (see Aghion and Howitt, 1998, pp. 109–10):

$$\dot{v}_t = [r_t + \phi_t]\, v_t - \pi_t \tag{20}$$

$$\dot{v}_{N,t} = [r_t + \phi_t]\, v_{N,t} - \pi_t \tag{21}$$

where π_t is the productivity adjusted profit of a local monopolist. It is worth recalling that, since an extensive innovation will be displaced by an intensive innovation in the same sector, the expected obsolescence rate takes the same value ϕ_t for extensive *and* intensive innovations.

Since the productivity adjusted value of extensive and intensive innovations are identical in equilibrium, $v_t = v_{N,t}$, which in steady state can be written:

$$(1+\delta)\lambda u_L^{\theta-1} m^{\theta} = \beta z_L^{\varepsilon-1} m^{\varepsilon} \tag{22}$$

Using (14) and (15) we obtain:

$$\frac{u_L}{z_L} = \left[(1+\delta)\lambda n^{\frac{\theta-\varepsilon}{\varepsilon}} \beta^{\frac{-\theta}{\varepsilon}} \right]^{\frac{1}{1-\theta}} \tag{23}$$

$$g_A = \delta \left[\lambda(1+\delta)^{\theta} n^{1-\varepsilon} \beta^{\frac{-\theta}{\varepsilon}} \right]^{\frac{1}{1-\theta}} \tag{24}$$

In the special, but convenient case $\theta = \varepsilon$ (23) and (24) simplify to:

$$\frac{u_L}{z_L} = \left[(1+\delta)\lambda\beta^{-1} \right]^{\frac{1}{1-\theta}} \tag{23'}$$

$$g_A = \delta n \left[\lambda(1+\delta)^{\theta} \beta^{-1} \right]^{\frac{1}{1-\theta}} \tag{24'}$$

Thus we reach the striking conclusion that in the endogenous model as specified above, an identical marginal innovation cost for intensive and extensive R&D makes (u_L/z_L) and g_A depend only on technological parameters. Instead, the steady-state shares u_L, z_L, and h_L depend also on the preference parameters ρ and σ. (See footnote 14, which refers to the special case $\theta = \varepsilon$.)

The reason why the model still qualifies as endogenous is that a policy variable such as an innovation subsidy (see Aghion and Howitt, 1998, p. 419) would affect the rate of growth if it exerted an asymmetric influence on the cost from one additional unit of labour effort in extensive and intensive R&D. For a discussion of this point, the reader is referred to the case considered in appendix A, where the cost asymmetry does not arise from a policy variable, but from a slight generalisation of the innovation technology considered above.

12.3.2. Non-endogenous R&D Growth

Referring back again to system (13), the crucial assumption of the non-endogenous R&D growth models is Det$[I - \Gamma] \neq 0$. In particular, referring to the case $[I - \Gamma]^{-1} > 0$, standard theorems of linear algebra lead to the following proposition, which shows that the result similar in spirit to be found in Eicher and Turnovsky (1999a) extends to our economy with expanding varieties and technology levels.

Proposition 12.3.2.1: Assume $\Gamma \geq 0$. Assume also that, for each row, the row sum of the elements of Γ is positive and lower than 1. Then, for every $n > 0$, there exist positive values g_A, g_N, g_K that are solutions to (10), (11) and (12) and such that $g_l = n - g_N > 0$.
Recalling that $0 < \alpha < 1$, a quick look at equation (12) reveals:

Proposition 12.3.2.2: If, in addition to the assumptions of proposition 12.3.2.1, we have $(\gamma + 1 - \alpha) \geq 0$, then $g_K > n$ (positive per-capita-output growth).

Remark 12.3.2.1: The *if* condition of Proposition 12.3.2.2 amounts to the existence of increasing returns to scale in the output sector. The assumption of Proposition 12.3.2.1 implies, but is not equivalent to, aggregate decreasing returns to scale in extensive and intensive search.

Thus, where the equations of system (13) are not linearly dependent (notably, a condition of full measure in the relevant parameter space) the steady-state growth rates of output, technology levels and varieties are completely determined by population growth and the technological parameters. These rates are therefore independent of preferences and of savings rates in particular.

The above propositions extend to a three-sector environment the formal characterisation of the class of two-sector non-endogenous growth models first laid down by Eicher and Turnovsky (1999a). From a formal viewpoint the seminal paper of Arrow (1962), where technology accumulation is driven by learning rather than deliberate R&D investment, belongs to the same class. Within the family of R&D growth models, the best-known non-endogenous example is probably that provided by Jones (1995; see also Jones, 1998 and 2002), where the author abstracts from the expansion of varieties, so that $g_N = 0$ and $g_I = n > 0$. In particular, Jones (1995) assumes $\xi = 0$ (no physical capital input in R&D) and $0 < -\chi < 1$, so that his two-sector version of system (13) boils down to

$$\begin{bmatrix} -\chi & 0 \\ -(1-\alpha) & (1-\alpha) \end{bmatrix} \begin{bmatrix} g_A \\ g_K \end{bmatrix} = \begin{bmatrix} \theta n \\ (1-\alpha)n \end{bmatrix}$$

and the conditions of propositions 12.3.2.1 and 12.3.2.2 are trivially satisfied.

Interestingly, the steady-state relation $g_c = g_A = (r - \rho) / \sigma$ continues to hold, but the direction of causality at work here is such that, given n, technology determines g_A and r is then determined by g_A and preferences. As is discussed in appendix A, in the endogenous model with asymmetric cost of innovation effort between extensive and intensive R&D, technology and preferences simultaneously determine g_A and r.

12.4. IS n AN UPPER BOUND FOR g_N?

As it turns out, the available examples of endogenous and non-endogenous R&D growth models share the prediction that, *in steady state*, the expansion of varieties proceeds at a pace which is *not faster* than the pace of population growth. In particular, $g_N = n$ in the endogenous and $g_N < n$ in the non-endogenous models considered above. On closer examination, however, these predictions are the by-product of quite special assumptions. Both the endogenous and the non-endogenous model admit extensions such that g_N may be greater than n. The point is considered in appendix B.

12.5. RESEARCH EMPLOYMENT AND PRODUCTIVITY

A second and deeper problem is posed to the R&D growth models by the stylised facts *(a)* and *(b)* mentioned in the introduction. These stylised facts are at variance with the possibility of approximating (if at a very aggregate level) the long term evolution of innovation activity and productivity growth in the U.S. through the hypothesis that this economy has been growing in the neighbourhood of a single steady-state path. More specifically, endogenous and non-endogenous models alike are faced with the problem of

(*i*) explaining how the rising researchers/employment ratio $(1 - h_L)$ can be reconciled with the behaviour of productivity growth;
(*ii*) identifying the causes of the rising researchers/employment ratio.

A first way of answering these questions is to suppose that the rise in $(1 - h_L)$ corresponds to a transitions path with constant growth rate g_A induced by *exogenous* changes in one or more technological parameters.

A second and more ambitious way is much in the spirit of Pasinetti (1981) and searches for *rules of structural change* that may get closer to explaining the observed phenomena without resorting to exogenous parameter changes. In the remainder of this paper we shall expand on these two lines of investigation.

To this end, I shall refer to the simplified versions of system (13) that feature in 'standard examples' of endogenous and non-endogenous R&D growth models. In particular, physical capital is not an input to innovation activity, intensive and extensive, hence $\xi = 0$, $\psi = 0$; the productivity of the extensive innovation effort does not depend on the technology level A, that is, $\nu = 0$; the aggregate production function does not depend on the number of varieties N, thus $\gamma = \alpha - 1$. In addition, I introduce the simplifying restriction $\varepsilon = \theta$, that is, the elasticity of innovation output with respect to R&D labour effort is uniform across extensive and intensive innovations.

12.5.1. Looking for Appropriate Parameter Changes

Referring to the U. S. experience in the second half of the twentieth century, we may observe how the rate of interest and the capital output ratio have been 'relatively constant'[11] over the period. Since the model structure implies $\sigma g_A + \rho = r = \alpha^2 K / Y$, using stylised fact (*b*) we derive the restriction that α has been constant; in this Section I am also led to formulate the 'working hypothesis' that the preference parameters σ and ρ were unchanged throughout. With this situation in mind I consider what, if any, changes of the technological parameters of the non endogenous and endogenous models can answer the issues posed under (*i*) and (*ii*) above.

12.5.1.1. Non-endogenous model

With the assumptions of proposition 12.3.2.1 in place, in particular $0 < -\chi < 1$, $\varepsilon + \tau < 1$, the non-endogenous model yields the steady-state predictions:

$$g_Y = g_A + n$$

$$g_N = \frac{\varepsilon n}{(1-\tau)}$$

$$g_A = \frac{\theta(1-\tau-\varepsilon)n}{-\chi(1-\tau)}$$

The growth rate of per capita output is independent of δ, the proportional productivity effect of quality innovations; it is also independent of λ and β, the parameters that, for any given innovation effort, regulate the arrival rates of intensive and extensive innovations, respectively.

Since the (expected) productivity-adjusted values v_t, $v_{N,t}$ of intensive and extensive innovations are identical, free entry in R&D implies the following equilibrium condition at every date t:

$$\frac{u_{L,t}}{z_{L,t}} = \left[A^{\chi}_{t,Max}(1+\delta)\frac{\lambda}{\beta} N_t^{1-\tau-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \tag{25}$$

On every equilibrium path sustained by smooth changes of λ, β and δ:

$$g_{A,t} = \left[\frac{\dot{\delta}_t}{1+\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} - \frac{\dot{\beta}_t}{\beta_t} + (\theta-1)\left(\frac{\dot{u}_{L,t}}{u_{L,t}} - \frac{\dot{z}_{L,t}}{z_{L,t}} \right) + (1-\tau-\varepsilon) g_{N,t} \right] \frac{1}{-\chi} \tag{26}$$

On a growth path with constant growth rates of A and N:

$$g_{N,t} \left[\frac{\dot{\beta}_t}{\beta_t} + \theta\left(\frac{\dot{z}_{L,t}}{z_{L,t}} + n \right) \right] (1-\tau)^{-1} \tag{27}$$

$$g_{A,t} = \left[\frac{\dot{\delta}_t}{\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} + \theta\left(\frac{\dot{u}_{L,t}}{u_{L,t}} - \frac{\theta}{1-\tau}\frac{\dot{z}_{L,t}}{z_{L,t}} - \frac{1}{1-\tau}\frac{\dot{\beta}_t}{\beta_t} + \frac{1-\tau-\varepsilon}{1-\tau} n \right) \right] \frac{1}{-\chi}$$

Substituting from (27) into (26) we obtain:

$$-\frac{\dot{\delta}_t}{\delta(1+\delta_t)} = \frac{\dot{u}_{L,t}}{u_{L,t}} - \frac{\dot{z}_{L,t}}{z_{L,t}} \tag{28}$$

Using (28), (27) and (25), a transition path with

$$\frac{\dot{\delta}_t}{\delta_t} \neq 0, \frac{\dot{\lambda}_t}{\lambda_t} \neq 0, \frac{\dot{\beta}_t}{\beta_t} \neq 0$$

and constant growth rates $g_{A,t}$, $g_{N,t}$ satisfies:

$$\text{if } \frac{\dot{\delta}_t}{\delta_t} < 0 \text{, then } \frac{\dot{\delta}_t}{\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} > \frac{\dot{\beta}_t}{\beta_t} \quad ; \quad \text{if } \frac{\dot{\delta}_t}{\delta_t} > 0 \text{, then } \frac{\dot{\delta}_t}{\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} < \frac{\dot{\beta}_t}{\beta_t}.$$

Moreover, the steady state share $(u_L + z_L)$ is independent of λ and β and satisfies[12] $\partial(u_L + z_L)/\partial\delta < 0$, if σ is not too lower than 1. The above considerations suggest the conjecture that a transition path with rising share $(u_L + z_L)$ and constant growth rate of productivity is explained by

$$\frac{\dot{\delta}_t}{\delta_t} < 0 \text{ and } \frac{\dot{\delta}_t}{\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} > \frac{\dot{\beta}_t}{\beta_t}.$$

To gain some understanding of the problems posed by this lone of reasoning, it is worth observing that the dramatic rise of $(u_L + z_L)$ would be obtained through partly offsetting changes of u_L and z_L. Recalling (28), our conclusion here is that the rates of change of the technological parameters δ and λ which are required to explain the stylised facts *(a)* and *(b)* may be implausibly high.

12.5.1.2. Endogenous model

In addition to the simplifying assumptions stated at the outset of section 12.5, the endogenous model we are considering assumes $\chi = 0$, $\varepsilon + \tau = 1$. The innovation technology is that considered in section 12.3.1.1 generating a symmetric cost from one additional unit of labour effort across extensive and intensive innovations.[13] Following the same line of reasoning explained above, we obtain that a transition path with smooth changes of λ, β and δ and constant growth rates $g_{A,t}$, $g_{N,t}$ satisfies (28) and

$$\frac{\dot{\delta}_t}{\delta_t} + \frac{\dot{\lambda}_t}{\lambda_t} = \frac{\dot{\beta}_t}{\beta_t} + \delta_t\left(\frac{\varepsilon}{\delta_t(1+\delta_t)} \right).$$

The difficulties encountered by the line of reasoning under investigation are therefore similar to those discussed for the non-endogenous model.

12.5.2. Growth and Structural Change

In a recent paper, Kongsamut, Rebelo and Xie (2001) suggest that the long term rise in the service-employment share has to do with changes in the composition of consumers' expenditure associated with the long term rise of per-capita income. A tradition in economic theory, from Kuznets (1957) to Pasinetti (1981) had already emphasized this order of phenomena. In a similar vein, I introduce in this Section the hypothesis that the long-term rise of the research employment share may be explained by a slow, almost negligible secular rise of the intertemporal elasticity of substitution associated to the long-term rise of GDP per capita[14]. In what follows, the focus of my analysis is not that of giving a detailed specification of the hypothesis, but is that of suggesting a line of argument explaining how stylised facts *(a)* and *(b)* may be reconciled.

The explanation rests upon the complementarity between the goods and methods used in production. However often neglected, the idea is far from new. Perceptive remarks on the relevance of this notion can be found in Marx's volume I of *Capital*. In chapter XV it is emphasised that the successful exploitation of new engineering and scientific principles in production required the emancipation of technology from the pre-existing set of tools[15] and that 'a radical change in the mode of production in one sphere of industry involves a similar change in other spheres' (Marx, 1887, p. 362). In the new-growth literature, the problem of complementarity between intermediate goods has been introduced in relation to the idea of a sequence of general-purpose technologies (GPTs). The adoption of a GPT requires the previous creation of a set of intermediate goods that are specific to it.[16]

I suggest that a similar set of ideas can be conducive to phenomena of structural change within a framework which is borrowed, with some important variations or qualifications, from the R&D growth models considered in this paper.

For the sake of simplicity, let us assume away the problem of extensive R&D by assuming that at every date there is an unchanging continuum of intermediate-good varieties ordered on $\Re_+$. To employ these varieties in production, their appropriate technology level must be developed. $[0, \Lambda_A]$ is the set of complementary intermediate-good inputs *necessary* to implement the technology level A in the production of final output. N_t is the number of intermediate goods *used* at t. There is only one final good Y. Its production function is:

$$Y_t = N_t^{\alpha-1} L_{Y,t}^{1-\alpha} \left[\int_{v=0}^{\infty} P(A_{v,t}) x_{v,t}^{\alpha} \partial v \right]$$

where $P(A_{v,t})$ is the productivity index associated to the technology level $A_{v,t}$ of variety v. with $P(A_{v,t}) = A$, if $0 \le v \le \Lambda_A$ and $A_{v,t} = A_{j,t} = A$ for all $v, j \in [0, \Lambda_A]$; $P(A_{v,t}) = 0$ otherwise. This assumption formalises a strong form of incompatibility between intermediate goods of a different technology level.

We say that technology level A has been implemented if $A_{v,t} = A_{j,t} = A$ for all $v, j \in [0, \Lambda_A]$ .Variety v is *necessary* to the implementation of A if and only if $\in [0, \Lambda_A]$.

If technology level $A(t)$ is implemented at time t, there is an instantaneous knowledge spillover such that $A_{v,t} = A(t)$ for every $v \in [0, \infty]$. The implementation of a *higher* technology level is instead costly, because it requires the higher level is independently developed for every necessary variety as the result of a deliberate R&D effort. The number $\phi_{v,t}$ of intensive innovations in sector v at t evolves according to the *deterministic* process:

$$\phi_{v,t} = \lambda (u_{L,v,t} L_t)^{\theta} A_{v,t}^{\chi}$$

If every innovation has a proportional effect δ on the technology level $A_{v,t}$, we obtain:

$$\dot{A}_{v,t} = \delta\lambda \left(u_{L,t} \frac{L_t}{N_t} \right)^{\theta} A_{v,t}^{\chi+1} \tag{29}$$

Higher technology levels are of higher complexity and their implementation requires a larger number of necessary intermediate inputs. Assume that the number of necessary varieties evolves according to:

$$\Lambda_{A(t)} = A_t^{\eta} \qquad \eta > 0$$

This implies that, if $g_{\Lambda(t)}$ is the proportional growth rate of $\Lambda_{A(t)}$, then:

$$g_{\Lambda(t)} = \eta\, g_{A(t)} \tag{30}$$

The strong complementarities of the form described above imply that the market implementation of a higher technology level will face a host of co-ordination problems. Here we are not concerned with this feature, however important it may be. Our aim is simply to show that equilibrium paths on which the productivity index A_t grows at a positive constant rate $g_A > 0$ are *not* steady states and have a rising share u_L of R&D employment.

In equilibrium, $N_t = \Lambda_{A(t)}$. With g_A constant, from (29) and (30) we obtain:

$$-\chi g_A = \theta \left(n + \frac{\dot{u}_{L,t}}{u_{L,t}} - g_{\Lambda}(t) \right)$$

$$\left[\eta - \frac{\chi}{\theta} \right] g_A = \left(n + \frac{\dot{u}_{L,t}}{u_{L,t}} \right)$$

Recalling that the 'congestion effect' in R&D implies $\theta < 1$, and that our considerations suggest $\chi < 0$, it is easy to see that, given n, the higher the value of η, the higher the growth rate $\dot{u}_{L,t}/u_{L,t}$ required to elicit a given productivity growth g_A . Thus, with η sufficiently large, the value $g_A \approx 0.02$ prevailing in the period 1950–93 would not have been possible in the presence of a constant labour share in R&D. Indeed, a growth rate g_A of the observed dimension cannot be a steady-state growth rate and cannot be sustained 'for ever'.

If the argument above offers a tentative explanation of how the long-term rise of the researchers/employment ratio can be reconciled with a constant growth rate of productivity, what is yet to be explained is the source of the rising researchers/employment ratio.

Here I suggest, as a working hypothesis to be explored by future work, that the preference structure with constant inter-temporal elasticity of substitution is replaced by a preference structure such that the rising per-capita consumption causes a slowly rising inter-temporal elasticity of substitution. Since h_L is close to 1 and u_L is close to zero,[17] the required change in σ does not have to be large, since a very small, seemingly negligible, shift away from employment in manufacturing in favour of research is sufficient to explain that:

(*i*) $\dot{h}_{L,t}/h_{L,t}$ is negative but very close to zero, as in the data;

(*ii*) $\dot{u}_{L,t}/u_{L,t}$ is positive and significantly large, as in the data.

12.6. CONCLUSIONS

In spite of the statements to the contrary (Jones, 1995; Aghion and Howitt, 1998, ch. 12), growth models that avoid the scale effect of R&D employment on productivity growth do not explain the evidence on R&D employment and productivity growth in the U.S. Indeed, the stylised facts (*a*) and (*b*) mentioned in the introduction are not easily reconciled within the standard steady-state hypothesis.

The first reason offered in this paper is that cross-sector research spillovers are less extensive than is normally assumed in R&D models: After a new basic idea is first discovered, the development and profitable implementation of the same idea in the production of final utput is a costly process. A second reason is that the number of complementary inputs necessary to implementa technology level A in the production of final output is likely to be an increasing function of A. The further assumption of complementarities in the form of strong incompatibilities between intermediate goods of a different technology level yields the result that structural change in the form of a rising R&D employment share is a necessary condition for the sustained growth rate of productivity experienced in the U.S. during the second half of the twentieth century.

APPENDIX A: ASYMMETRIC INNOVATION COST

Suppose that every unit of labour invested in R&D at time t is combined with a quantity of capital $A_{t,\mathrm{Max}}\, T_A$, in the case of intensive R&D and $A_t\, T_N$ in the case of extensive R&D. In this section I assume $T_N \neq T_A$. In other words, labour and capital are perfectly complementary inputs to intensive and extensive innovation activities, but the ratio between the two inputs is different in the two sets of activities, even after adjustment is made for the productivity levels $A_{t,\mathrm{Max}}$ and A_t. The case $T_N = T_A$ yields conditions identical to those obtained in Section 12.3.1.1, with the understanding that terms K and q must be replaced everywhere with $h_K K$ and $h_K q$, where h_K is the fraction of total capital employed in the output sector (to produce intermediate goods). u_K and z_K are the fractions of total capital employed in intensive and extensive R&D, respectively. With this notation, and assuming for simplicity $\theta = \varepsilon$, the procedure followed in Section 12.3.1.1 yields:

$$\frac{u_L}{z_L} = \left\{ \frac{\lambda(1+\delta)\left[(1-\alpha)+\alpha^2\left(\dfrac{r}{\alpha^2}\right)^{\frac{1}{1-\alpha}} T_N\right]}{\beta\left[(1-\alpha)+\alpha^2\left(\dfrac{r}{\alpha^2}\right)^{\frac{1}{1-\alpha}} T_A\right]} \right\}^{\frac{1}{1-\theta}}$$

Hence u_L/z_L is related to the steady-state rate of interest, which depends on the preference parameters ρ and σ. In particular, it can be easily checked that the sign of $\partial(u_L/z_L)/\partial r$ is positive if $T_N - T_A > 0$ and is negative if $T_N - T_A < 0$. Moreover, similar considerations apply to the relation between g_A and the rate of interest. We can write:

$$g_A = \frac{r-\rho}{\sigma} = f(r, \lambda, \delta, \beta, \alpha, \theta, T_N, T_A)$$

If $T_N - T_A \neq 0$, then r is a non-redundant argument of the function $f()$ and, given n, g_A and r are simultaneously determined by technology and preferences. If $T_N = T_A$ the simultaneity collapses and g_A is determined by (24′).

APPENDIX B: EXISTENCE OF SOLUTIONS WITH $g_N > n$

It is enough to provide two examples: one for the endogenous model and one for the non-endogenous model. In both examples the simplifying restriction $\gamma = \alpha - 1$ holds so that Det$[I - \Gamma] = -(\xi + \chi)(1 - \tau - \varepsilon - \psi)(1 - \alpha)$.

For the endogenous model with $\chi = 0$, $\xi > 0$, $\upsilon > 0$, the crucial restriction Det$[I - \Gamma] = 0$ is fulfilled by $\tau + \varepsilon + \psi = 1$. In this case

$$\dot{N} = N\beta z_L^{\varepsilon}(L/N)^{\varepsilon} A^{\upsilon}$$

which in steady state requires $\varepsilon(n - g_N) + \upsilon g_A = 0$. If $0 < \upsilon < \varepsilon$, this yields $g_{Ki} = g_A + n > g_N$. Since from (10) $g_K = g_N - (\theta / \xi)(n - g_N)$ we have that $g_N > n$ and $g_A > 0$ are consistent with a steady state path.

For the non-endogenous model it is sufficient to assume $\tau < 1$, $\tau + \varepsilon + \psi > 1$; $\xi + \chi > 0$, ν and ψ sufficiently close to zero.

NOTES

1. There are instances of R&D activities performed in a given country which exert their productivity effects mainly outside the country: think of a new treatment for curing a tropical disease discovered in the U.S.A. or in Germany. The view taken in this paper is that this type of phenomenon is far from explaining the qualitative evidence presented in the text. I thank Francesco Pigliaru for drawing my attention to this point.
2. To reconcile facts (*a*) and (*b*), two candidates come to mind. (*i*) There has been a fall in the average effect of innovations on measured productivity. This may be at least partly due to the fact that official statistics underrate the qualitative changes in goods and the improvement in their service characteristics (Nordhaus, 1997). Alternatively, or in addition to the previous cause, it may be increasingly difficult to produce the same proportional improvement in the service characteristics of goods. Hence, the productivity gain tends to fall in the more recent innovations. Robert Gordon (2000) compares the effects on well-being of the 'new economy' to those produced by the great innovations during the second industrial revolution. He concludes that the effects of the former do not bear comparison with those of the latter. (*ii*) A different, but compatible, line of explanation is a fall in the average productivity of R&D labour, as measured by the number of innovations per unit of research effort. A fall of this kind has certainly taken place, if the number of innovations is measured through the number of patents, granted or applied for (Griliches, 1988, 1990). Measures of this type are strongly biased not only by changes in the 'productive capacity' of institutional patent agencies (e.g. the U.S. Patent Office), but also by changes in the propensity to apply for a patent. Microeconomic studies (Lanjow and Schankermann, 1999) indicate that a lower fall of the productivity of R&D labour is obtained if the aggregate innovation output is obtained by weighting patents by means of indicators of their technological and economic importance. This is related to point (*i*) above.
3. We shall not consider other families of models where growth is likewise driven by innovations, let alone the huge microeconomic literature on R&D.

4. By definition, on a steady-state path the growth rate of every variable is constant for ever. Since the employment shares are bounded between zero and one, their unique admissible steady-state growth rate is zero.
5. Peretto (1998) reports on the transition dynamics of an R&D growth model where the endogenously accumulating factors are only A and N. In the transition dynamics results of Aghion and Howitt (1998, pp. 109–15), the endogenous factors are A and K.
6. Still, in reading it, it is best to bear in mind what is implied by the seminal work by Jacob Schmookler on innovation and growth: the interest in the causes of the long-term growth of GDP per capita, as distinguished from the GDP level, is at best only a partial justification for the rigid supply orientation of general-equilibrium growth models.
7. The assumption is not fully realistic. Even granting that A_v amounts to a productivity index, we should in general expect the flow of service characteristics associated with (v, A_v) to depend upon the type and quantity of other intermediate goods with which (v, A_v) co-operates within a production activity. If there are strong complementarities between different intermediate goods, the best-practice technology level of variety v at t may not be the highest available. Compatibility constraints may in fact imply that it is inefficient to use very different technology levels of complementary varieties in the same activity. Complementarities of this sort are simply ruled out in most (an exception is Helpman and Trajtenberg (1994); see section 12.5.2 below) R&D growth models.
8. If $1 > \omega$, then the monopoly output is positively related to the technological advance $A_{v,t}$.
9. Realistic as it may be, the positive correlation between the technology-frontier index and the difficulty of search must be simply assumed and cannot find a micro foundation within a formal framework which does not lend itself to consider the feedback of innovations on the complexity of the search space.
10. As before, since the present framework cancels from view the rising complexity of the technology space, the treatment of this feature can be at best evocative.
11. At least in the sense specified in the introduction to this paper.
12. $(u_L + z_L)^{-1} = 1 + \dfrac{\theta n(1-\tau-\varepsilon)(1+\sigma\delta) - \chi\delta[(\rho-n)(1-\tau-\varepsilon)+\rho\varepsilon]}{(1+\delta)(1-\tau-\varepsilon)\alpha\theta n - \chi\delta\varepsilon\alpha n}$
13. The research employment shares are:

$$u_L = \frac{\left[(1+\delta)\dfrac{\lambda}{\beta}\right]^{\frac{1}{1-\varepsilon}}}{1+\dfrac{\rho}{\alpha n}\left[1+\dfrac{\sigma\delta+1}{(1+\delta)\alpha}\right]\left[(1+\delta)\dfrac{\lambda}{\beta}\right]^{\frac{1}{1-\varepsilon}}}$$

$$z_L = \left[1+\frac{\rho}{\alpha n}\left[1+\frac{\sigma\delta+1}{(1+\delta)\alpha}\right]\left[(1+\delta)\frac{\lambda}{\beta}\right]^{\frac{1}{1-\varepsilon}}\right]^{-1}$$

It can be easily verified that : $\partial u_L / \partial(\lambda/\beta) > 0$; $\partial z_L / \partial(\lambda/\beta) < 0$; $\partial(u_L + z_L) / \partial(\lambda/\beta) < 0$ and $\partial(u_L + z_L) / \partial\delta < 0$ if $\sigma > [(1+\delta)\rho - n] / \delta n$; $\partial(u_L + z_L) / \partial(\lambda/\beta) > 0$ and $\partial(u_L + z_L) / \partial\delta > 0$ if $\sigma \leq 1$.
14. The hypothesis implicitly assumes some measurement error leading to a (very) mild under-evaluation of productivity growth. See above, note 2.

15. 'It is only after considerable development of the science of mechanics, and accumulated practical experience, that the form of a machine becomes settled entirely according to mechanical principles, and emancipated from the traditional form of the tool that gave rise to it' (Marx, 1887, p. 362, n. 1).
16. When the GPT s first appears a labour share is shifted from manufacturing to R&D (phase 1); next, after the intermediate goods required by s have been invented all employment is shifted to manufacturing until the GPT (s + 1) arrives (phase 2). The idea is exploited by Helpman and Trajtenberg (1994) and Aghion and Howitt (1998) to study the relation between growth and cycles. The notion of a steady state is correspondingly extended by these authors to the effect that in an economy with a constant population 'a steady-state equilibrium is one in which people choose to do the same amount of research each time the economy is in phase 1 …' (Aghion and Howitt, 1998, p. 248).
17. The U.S. researchers/employment ratio was 0.008 in 1993 (see Jones, 2000, p. 16).

13. Competition and technical change in Aghion & Howitt: a formalisation of Marx's ideas?

Maria Daniela Giammanco

13.1. INTRODUCTION

The aim of this paper is to identify some similarities between Aghion and Howitt's analysis and Marx's work with respect to competition and technical progress as presented in Aghion and Howitt's 1992 model, and some of the extensions proposed by the two authors in 1998.

The comparison is not based on analogies between the methods of analysis, which reflect two different visions of the world. Marx investigates the distribution of the surplus produced in the economy among conflicting social classes and conflicting capitals. Aghion and Howitt propose a neoclassical general equilibrium model whose parameters are individual initial endowments, individual preferences, technology and the probabilistic production function of new technology. The proposed comparison stems rather from the identification of some fundamental assumptions characterising Aghion and Howitt's model, which are also to be found at the basis of Marx's idea of competition and technical progress.

Aghion and Howitt's work belongs to the New Growth Theory (NGT) tradition. Romer (1994) affirms that the origins of the NGT must be sought not only in the effort to solve the 'convergence problem' but also in the effort to build a valid alternative to perfect competition at a theoretic level. The need to abandon perfect competition stems from the necessity to explain one of the determinants of economic growth ignored in the neoclassical model: endogenous technical progress. Aghion and Howitt (1992) acknowledge the pioneering contributions to the endogenous growth theory of Romer (1986c) and Lucas (1988) and propose a model in which economic growth is driven solely by innovation resulting from the firm's research activity. Such an activity is performed because, if it meets with success, monopoly rents are earned, though momentarily; these rents are possible because knowledge, as in Romer (1990b), not only has a non-rival component but is also a partially excludable good. Non-rivalry, together with partial excludability, introduces increasing returns internal to the firm and therefore imperfect competition.

An investigation of the characterising assumptions made by Aghion and Howitt will show that these assumptions are really close to the elements which characterise Marx's view of competition and technical change. A companion paper (Giammanco, 2002) focuses on the two conceptions of competition proposed by Marx. In Marx's analysis competition is a negative force and also a process. From the idea of competition as a negative force, destroying all barriers to the free development of capital follows a vision of technical progress as a powerful weapon of capital in its struggle for complete command over labour power, reduced to simple labour. Marx's idea of competition as a process, inseparable from the capitalistic mode of production where a complete command of capital over labour has already been established, leads to the idea of technical progress as a powerful tool in the struggle among capitals in the fully developed capitalistic society. Technical progress as an instrument in the struggle among capitals focuses on a major aspect of Marx's thought – the deliberate action of the capitalist/innovator moved by the desire/need to innovate, which engenders relentless technological change.

Aghion and Howitt do not refer to Marx as their *inspiring Muse*. However, their work, like many within the NGT tradition, is stimulated by the many empirical studies on technological change as a source of growth, engendered by profit-rent seeking firms. This is a cornerstone of Marx's vision; those who know Marx might find it stimulating to encounter this and some of his other crucial arguments treated formally within an elegant general equilibrium construction.

The first section of the work presents Aghion and Howitt's model of 1992; it also outlines the fundamental issues underlying their idea of competition and technical progress, which play a crucial role in Marx's analysis. The second section enlarges the analysis to some of extensions of the 1992 model proposed by Aghion and Howitt in 1998. The third section offers some conclusions.

13.2. AGHION AND HOWITT'S ANALYSIS

Aghion and Howitt (1992) focus their attention on technical progress improving the quality of products; by doing so, they consider the obsolescence created by technical progress which consequently engenders losses as well as gains. They refer specifically to Schumpeter's idea of

creative destruction and to the revolutionising of the economic structure by new production methods: hence they call their model neo-Schumpeterian.[1]

Aghion and Howitt present a general equilibrium model, with perfectly competitive final good and research sectors and a monopolised intermediate one. The final good is produced by the intermediate input together with the total endowment of unskilled labour of the economy whose amount is fixed. As a consequence the final good production function is represented by a constant returns function, the only argument of which is the intermediate good multiplied by a parameter representing the intermediate good's productivity and increasing by the same amount with every successful innovation. For each firm the invention of a new intermediate input is a random event governed by a Poisson distribution, with an arrival rate which is a constant returns to scale function, the arguments of which are researcher-specific skilled labour and specialised labour, each invention being an independent event at the aggregate level. The monopolistic intermediate sector produces the intermediate good using skilled labour together with a design: to find the equilibrium means to determine, in every period, the allocation of the given amount of skilled labour between intermediate and research activities.

The strongest analogy between Aghion and Howitt's work and Marx's analysis is that the mechanism generating growth in Aghion and Howitt's model is the mechanism which propels the accumulation of capital in Marx's analysis. In Aghion and Howitt's work monopolistic competition in the intermediate sector drives growth. In this sector, perfect competition is abandoned thanks to the introduction of the Schumpeterian idea of creative destruction. This idea is based on the assumption that knowledge is partially excludable: the firm that innovates obtains a patent that lasts forever, but will be used only till a better intermediate input is invented. Perfect competition cannot prevail in the intermediate sector, because the cost of the patent is a fixed cost which the innovating firm will pay only if it can sell its output at a price higher than the marginal cost. Technical progress by means of creative destruction makes the monopoly power of the successful innovating firm last only for a period, and creates new rent opportunities for the next innovator. Growth in Aghion and Howitt's model is therefore engendered by the prospect of monopoly rent which makes firms carry out research; it stems from rivalry among innovating firms. In Aghion and Howitt's analysis, the innovating firm plays the role of the capitalist/innovator whose action continuously revolutionises the economic structure in Marx's world. In both frameworks the existence of a patent system is crucial (on the patent system which serves to meet capitalists' demand for differentiation in Marx, see Giammanco, 2002). A major aspect of Marx's thought is in fact the deliberate action of the capitalist/innovator moved by the desire/need to innovate, which engenders relentless technological change; the introduction of a new technology within a productive sector raises labour productivity and allows for transient extra-profits, until it becomes widespread, igniting an endless struggle in the production arena. As profit rates tend to equalise, this attracts capitals towards the innovating sector and engenders ever-increasing demand for new and more efficient technologies (see Giammanco, 2002).

The strong analogy is therefore evident between Marx's analysis, in which continuous mutation in the economic structure is achieved by the incessant differentiation of each firm, and Aghion and Howitt's process of creative destruction, which engenders obsolescence and therefore winners and losers. However, while according to Marx firms show differences in the mode of exploiting common knowledge (see Giammanco, 2002), Aghion and Howitt, because of the general equilibrium framework chosen, do not formally tackle the problem of firm differentiation. In every period, the structure of the economy is characterised by perfect competition in all sectors but the intermediate one. Splitting their model into a perfectly competitive section and a monopolistic one, Aghion and Howitt try to confine the analysis of firm differentiation to the study of the monopolistic sector. In this sector, however, what changes from period to period is the firm which plays the role of the incumbent: while the present incumbent does not try to differentiate, outside firms strive to innovate and gain the monopoly position. This, however, does not engender differences in the research function available. The economic change results in changes in the productivity frontier of the final sector, which causes growth, and in a change of firm which assumes the role of incumbent in the intermediate sector.

Innovation as the mechanism generating growth may be further examined by looking at the backbone of Aghion and Howitt's model. The problem of finding the equilibrium is solved by satisfying an arbitrage condition in the labour market, together with a labour market clearing equation. The arbitrage condition establishes that for a positive amount of skilled labour used in research, the hourly wage paid in manufacturing must equal the expected value of an hour's research; the clearing equation states that the amount of skilled labour used in research and the amount used in manufacturing must add up to the given endowment of skilled labour. These two conditions embody the idea that at the basis of the functioning of the model is perfect competition, characterised by the free movement of assets towards the most remunerative use and by the frictionless nature of markets.

The arbitrage condition reflects the fact that labour can be freely employed either in the research sector or in the manufacturing sector, imposing that skilled labour be equally remunerated in both sectors. This condition stylises Marx's idea of the need for free mobility of capital towards the most remunerative uses: it states what competition as a negative force

should create, reducing labour to simple labour. By rendering the worker indifferent to the precise nature of the work, this facilitates inter-sectoral labour mobility – the crucial premise for capital mobility and the hallmark of a fully developed capitalistic society (see Giammanco, 2002). It must be said, however, that skilled labour in Aghion and Howitt's analysis cannot be considered as labour according to Marx, but as an asset which seeks the most remunerative use.

On the implication of the skilled labour market equation which reflects the frictionless nature of this market and assumes market clearing, it is worth referring to what Duménil and Lévi (1987) argue: the common characteristic of all dynamic equilibrium models consists in equality between supply and demand; it is possible, however, to conceive of different market mechanisms which respectively lead to: *i*) an *ex-ante* equilibrium in which demand immediately matches supply, as in the Walrasian model with production; *ii*) an *ex-post* equilibrium which is only reached asymptotically. The supply is determined together with the prices of the goods before the demand is known, and the information available is that of the previous period. The NGT uses the *ex-ante* equilibrium concept because it adopts an intertemporal equilibrium model, in which the Walrasian auctioneer determines the complete sequence of events in advance. Marx uses the *ex-post* equilibrium concept: the competitive process is characterised by adaptation to disequilibrium: the system gravitates around the equilibrium but does not reach it within the period considered.[2]

The interpretation of competitive equilibrium as an *ex-ante* or as *an ex-post* equilibrium modifies the rationale behind the mechanism generating growth. In Aghion and Howitt's model it is the amount of skilled labour devoted to research which determines growth. The average growth rate is determined through a forward-looking difference equation, which sets the amount of skilled labour employed today in terms of the amount of skilled labour employed tomorrow. What will be done tomorrow is known because of the perfect foresight assumption: perfect foresight is ruled out in Marx's analysis, in which technological progress is a path-dependent problem-solving activity (see Giammanco, 2002).

However, Grossman and Helpman (1994) argue that the NGT is well aware that firms struggle continuously at a micro level, and that growth is an irregular and stochastic process; but they argue that 'aggregation masks this micro-level turbulence' (p. 34). Ignoring the continuous change occurring in the competitive sector allows Aghion and Howitt to set the stage and give all the information necessary for the solution of the monopolistic maximization problem. This can be compared to the adoption of a long-run approach of classical inspiration,[3] to concentrate on the general laws governing the gravitational centres without investigating the analysis of disequilibrium which still remains at the heart of the competitive process.

Aghion and Howitt's choice to introduce a Poisson distribution regulating the outcome of research activity is a compromise between awareness of the uncertain outcome of research and the need to model through a well-defined probability function. Aghion and Howitt's concern shows close links with Marx's notion of the uncertainty of the innovative action. Marx speaks of the daring capitalist and is aware that trail-blazers often go bankrupt while money-capitalists, who merely acquire machinery and buildings at a low price from them, win the competitive struggle (see Giammanco, 2002). *Per se* the uncertainty characterising the outcome of innovative activity in Marx's analysis is not a statistically computable risk. However, a well-defined probabilistic distribution governing the outcome of research is an analytical necessity to model uncertainty within the limit imposed by a general equilibrium framework.

The aggregate production function of the final sector with its intermediate good productivity parameter introduces in Aghion and Howitt's analysis a scale effect that can be found, in slightly modified versions, in many New Growth Models (NGMs). Aghion and Howitt assume that the final output grows with the growth in the level of aggregate knowledge. Every successful innovation allows for an increase in the final sector productivity by a constant proportional factor, i.e., it allows for the use of more efficient methods in the production of the final good. The productivity rise due to successful innovation lasts forever, which introduces an important intertemporal spillover effect in the model. This scale effect in Aghion and Howitt's work implies that each new innovation allows other researchers to begin operating on the next one. At the macro level, size is therefore important. It is worth stressing that, though the intertemporal spillover effect is an aggregate phenomenon, it is limited to the final output sector, as the systemic accumulation of knowledge engenders aggregate increasing returns solely in the final output sector, while research activity is characterised by lack of memory (Aghion and Howitt, 1992, p. 327). This arises from the fact that the innovation arrival rate depends solely on present research and not on the stock of past research.

The scale effect in Aghion and Howitt's model recalls Marx's idea of the accumulation of experiences at a macro level, which become part of the general knowledge potentially available to everyone. In Marx the accumulation of general knowledge is one of the levers of the production of new ideas at the basis of technical progress. However in Marx's analysis the scale effect is both a micro and a macro phenomenon**;** *magnitude* is intimately connected with the capital accumulation process, being important both at the systemic and at the firm level. The systemic accumulation of

knowledge increases as the economy develops. As firms expand, the accumulated experience enlarges, with more possibilities for exploitation and for additions. There is also a strong link between collective labour, the possibility of exploiting ideas and scale economies: large plant-size allows not only constant capital saving, due to increasing returns to scale, but also the accumulation of practical experience, through the work of the combined collective worker, often necessary in order to exploit the common heritage of ideas (see Giammanco, 2002).

13.3. AGHION AND HOWITT (1998): SOME EXTENSIONS TO THE BASIC MODEL

Aghion and Howitt (1998, ch. 6) propose a further development to their 1992 model in which they try to represent the firm and its research effort. They acknowledge that in order to exploit research-generated knowledge, a firm must apply its own theoretical knowledge in practice, solve unexpected problems and grasp new opportunities. They therefore deal with heterogeneity in the innovative structure by distinguishing between fundamental and secondary innovations, which are complementary to each other. They consider two extreme and co-existing kinds of research: R&D and learning by doing.[4] R&D results in potential new products, whereas learning by doing results in improving existing ones. With learning by doing, Aghion and Howitt model an element which is at the basis of Marx's conception of technical change. Ideas can be produced not only by an R&D activity, but also as the outcome of a learning activity within production. Marx distinguishes between radical and incremental invention: a heritage of pre-existing ideas, *universal labour*, is a prerequisite for the birth of a new idea, thanks to which radical inventions may be produced. Each of these engenders a sequence of incremental inventions, thanks to the learning process of collective labour, and increases systemic knowledge (see Giammanco, 2002). Aghion and Howitt introduce the idea of general knowledge, 'the common scientific, technological, and cultural heritage potentially available to everyone' (p. 174), which is engendered by both R&D and learning by doing and which is really close to Marx's idea of universal labour.

Aghion and Howitt propose a first variant of the basic model, in which learning by doing is only accumulated at the macro level. They assume the existence of a constant mass of infinitely lived skilled labourers, each of whom can choose to work either in research or in productive sectors. There is an R&D sector which invents intermediate goods by means of research labour and general knowledge. Its outcome is ruled by a Poisson distribution.

The arguments of the aggregate production function of the final good sector are the quantity of labour used in the production of each intermediate good of different vintages, the quality of such intermediate goods, and the state of general knowledge. In this sector secondary innovations are jointly produced and are not internalised by each single firm: they improve the quality of already existing intermediate goods at the systemic level. The quality of the last vintage intermediate good is zero, continuously increased by the accumulation of systemic learning by doing. More recent vintage goods are potentially better, as they incorporate more general knowledge. Each firm producing an intermediate good of a specific vintage is a monopolist in that sector: it must compete with other firms in order to hire production labour.

As in the basic model, the steady-state growth rate is found by determining the steady-state proportions of workers engaged in research and in production. This is obtained by means of an arbitrage condition in the skilled labour market, determining the allocation of labour between research and production.

The importance of this variant of the basic model stems from the explicit introduction of learning by doing, which endorses the following ideas: innovations are not necessarily radical but can also be incremental; incremental technical progress can be a by-product of the production process which originates and develops new ideas aiming to solve practical problems; and the experience engendered by the production process accumulates at a systemic level. Aghion and Howitt make learning by doing appear indirectly in the aggregate production in two ways: by determining the quality level of the intermediate products of each vintage, and by influencing the level of general knowledge; the growth equation of general knowledge is in fact a function of the current flow of fundamental and secondary innovations and of the stock of previous general knowledge. The definition of general knowledge and the equation which governs its growth formalise Marx's conception of the systemic accumulation of knowledge: ideas and production experience become part of social knowledge accessible to everyone.

The production function of learning by doing, regulating the quality growth of intermediate goods of each vintage, not only models the accumulation of production experience at a macro level, which is strongly connected to Marx's learning process of combined collective labour which increases systemic knowledge; it also sketches the idea of the interdependence among productive sectors. This is so because the simultaneous amelioration in quality of all products can be explained by the close interdependence among sectors which forces technical progress to spread immediately. According to Marx, many of the productive sectors can be seen as gears of the same complicated mechanism. As a consequence, progress in one industry's mode of production revolutionises that of other industries (see Giammanco, 2002).

The interpretation of the arbitrage conditionis similar to that of the basic model. In this model, however, a fixed Poisson upgrading rate has been introduced, which is the rate at which a worker can switch from producing an old to producing a new vintage intermediate good, or to producing research.[5] From this it follows that only when skilled labour is free to move between research and the new vintage good sector, if it is exogenously upgraded, must its expected remuneration be equal in both sectors. This makes more explicit the importance of freedom of circulation of skilled labour, without which equality in remuneration ceases to exist, and creates a stronger link with Marx's analysis, according to which mobility of labour is a necessary condition for capital accumulation, hence growth.

Upgrading implies that the amount of production labour devoted to each intermediate good will fall exponentially with an increase in the age of the intermediate good: obsolescence is represented by a crowding-out process rather than a creative destruction process. Thus, Aghion and Howitt propose alternative dynamics to the process of creative destruction, interpreting capital struggle as a more gradual process, and take into account another aspect of technical progress present in Marx's analysis: Marx believes that technological trajectories have unknown potentialities which are discovered only with the passing of time (see Giammanco, 2002).

Aghion and Howitt introduce the idea of product life-cycle: as old intermediate goods incorporate less aggregate knowledge than new ones, with the introduction of new intermediate goods the contribution to production of old vintage goods decreases continuously as they become obsolete. This creates another analogy with Marx, who envisages a sort of product life-cycle for each new machine, a prototype that can be improved in the production phase (see Giammanco, 2002).

Aghion and Howitt also propose a further variant of the basic 1992 model, in which learning by doing is completely internalised by the firm.[6] Only the firm that solves problems within the productive process is able to improve its product quality. This means that as production workers can appropriate some of the fruits of their learning-by-doing activity, they must be compensated accordingly in a competitive equilibrium. The introduction of internalised learning by doing formalises another important element of Marx's analysis: the idea of learning through the accumulation of practical experience by the combined worker, which increases labour productivity within each firm.

With internalised learning by doing, Aghion and Howitt find that the size of the adaptability parameter is positively related to the growth rate, as long as the growth of fundamental knowledge does not depend too much on learning by doing. They explain this as a consequence of the increase in the share of skilled labour devoted to a new vintage good, which increases research productivity and reduces the cost of implementing an innovation.[7] This result further stresses the importance of worker adaptability in Aghion and Howitt's model[8] and its analogy with the importance of labour flexibility, stemming from the action of competition as a negative force, in Marx's work.

Aghion and Howitt also study another fundamental/secondary dichotomy by focusing on research and development. Research produces fundamental innovations, each consisting of a potential line of new products; development produces a workable plan for actually producing one such potential product. They explicitly model Marx's ideas on the difference between radical and incremental inventions. The method of analysis is analogous to that applied to learning by doing: also in this case there is a positive relation between growth and the workers' adaptability, provided that the growth of fundamental knowledge does not depend too much on development.

Aghion and Howitt (1998, ch. 7) consider not only product market competition but also competition as increased freedom of entry in the research sector. They identify the competitiveness level with the entry cost, and demonstrate that a higher degree of competitiveness in the research sector stimulates growth because it stimulates innovation. They refer to their basic model of 1992 and relax the assumption of constant returns to scale in the research sector; this is still governed by a Poisson arrival rate, defined as an increasing and concave function of the total labour employed in the industry minus the entry cost, multiplied research productivity. This allows them to find, as in the basic model, the steady-state growth rate which responds positively to an increase in competitiveness, i.e. to a reduction in the entry cost. According to Aghion and Howitt this result vindicates Schumpeter's claim that the higher the competition in research, the higher the growth. Also in Marx's analysis the presence of barriers to entry in the more technologically advanced sectors is an element restricting the exploitation of technological advances which are potentially available to everyone (see Giammanco, 2002). Accordingly, Aghion and Howitt, dealing with barriers to entry in the research sector, model another crucial idea which characterises Marx's analysis of capital accumulation.

13.4. CONCLUSION

In the previous pages I have suggested that the ideas underlying Aghion and Howitt's 1992 model, and part of its extension of 1998, can be found in Marx's analysis of competition and technical change. The analysis shows that in Marx, competition as a process is identified as the capitalistic mode of

production itself and is characterised by the struggle among capitals. In this struggle, technical progress is a very powerful weapon: it is the outcome of deliberate research activity and allows the innovator/capitalist to increase his profit. These ideas are developed by Aghion and Howitt's 1992 analysis, which deals with creative destruction. Marx's systemic accumulation of knowledge can be considered modelled, in Aghion and Howitt (1992), by the intertemporal spillover effect which increases productivity at the systemic level.

Aghion and Howitt's 1998 extension, in which they introduce two forms of learning by doing, matches an important notion of Marx's analysis: the difference between radical and incremental innovation. Learning by doing accumulated at a macro level also models the idea of Marx's systemic accumulation of knowledge development, by means of accumulation of experiences at a macro level. Learning by doing, increasing the productivity of already existing intermediate goods, also shapes the interdependence in technological advancement of different sectors of production present in Marx's analysis. The learning by doing internal to each firm recalls the importance of the learning process by means of Marx's combined worker. The relevance of the worker's adaptability parameter models Marx's idea of the importance of the extent of capital/labour mobility, which lies at the basis of the action of competition as a 'negative force'. The product life-cycle is also a problem found in Marx. The 1998 extension, in which Aghion and Howitt introduce barriers to entrance, models Marx's idea of a minimum plant-size required to implement inventions.

While there are ideas characterising Marx's treatment of competition and technical change that underlie Aghion and Howitt's analysis, attention has also focused on the parts of Marx's analysis that have not been developed by these authors. Although in their models competition drives firms to innovate in order to gain monopoly power, their analysis greatly simplifies the representation of the struggle for diversity characterising Marx's work. Moreover, Aghion and Howitt (1992) are aware of the uncertainty of the innovation activity considered by Marx and they introduce its proxy, which is risk: the invention of a new intermediate input is a random event governed by a single Poisson distribution. This models the firm's effort to innovate, hence to differentiate, but does not take into account that each firm, through research, will differ from other firms and have a distinct research production function. The same can be said of the 1998 extension which, with the entry barriers hypothesis, envisages identical firms facing the same entry cost and the same functional form of the arrival rate. These hypotheses are related to Aghion and Howitt's reliance on a general economic equilibrium construction, typical of the NGMs based on agents' maximization, which as argued by Nelson and Winter (1974) does not lend itself to deal with firms' differences in technology, profitability, knowledge and luck. In equilibrium no firm modifies its strategy.

In conclusion, it may be of interest to know that some of the best contemporary literature on growth develops in a rigorous and elegant way many ideas that loom large in Marx. The presence of many of Marx's ideas in Aghion and Howitt's model proves the sharpness and modernity of Marx's vision of the capitalistic society. Although Aghion and Howitt's formalisation can offer a clear view of some of Marx's arguments, the two authors do not tackle other crucial features present in his work. Marx observes history and proposes his explanation of the development of the capitalist society, the multiform aspects of which cannot be bridled by the general equilibrium approach. His analysis can still be a great source of stimuli for the growth theorist, who should not abandon formalisation but look for a more suitable analytical apparatus, which as suggested by Nelson (1998), should be more attentive to Knightian uncertainty and firm organisation.

NOTES

1. Romer (1994) defines as neo-Schumpeterian all the New Growth Models that abandon the hypothesis of perfect competition and whose discoveries are the outcome of a monopoly profit-seeking activity.
2. On competition as a tendency toward a predicted result, already present in Smith, see McNulty (1967 and 1968).
3. Duménil and Lévi (1985) demonstrate that, in the intertemporal equilibrium model with infinite-time horizon, Walrasian prices are the same as production prices only in an asymptotic position. Production prices are the prices depending on technology and distribution and independent of initial endowments and utility functions, i.e. of supply and demand. This is in line with what Kurz and Salvadori ((1998b and 1999) argue on the equivalence between NGT and the classical authors in the long run.
4. The implication of this model is that there exists a value of the growth rate beyond which an increase in research, at the expense of secondary innovations, jeopardises growth.
5. To shift from the production of an old vintage good to a new one is upgrading because the present value of the wage received by a worker who produces a good of vintage τ is a constant value multiplied by the general state of knowledge.
6. The implication of this extension is that when firms do not internalise learning by doing it is impossible for the level of research to be too high.
7. This effect, by means of time discounting, dominates the reduction in the profitability of research due to the consequent increase in the rate at which a worker will quit the newly discovered line of production.
8. Aghion and Howitt consider the positive relation between growth and adaptability the result of an increase in the amount of research labour, and not, as in Lucas (1993), the result of enhancing learning by doing.

14. Division of labour and economic growth: Paul Romer's contribution in an historical perspective

Andrea Lavezzi

14.1. INTRODUCTION

The issue of economic growth regained importance in the economic profession with the development of the endogenous growth theory (EGT).

Among the plethora of models that followed the seminal papers of Romer (1986b) and Lucas (1988), an old hypothesis re-emerged: that economic growth depends on increasing returns generated by the division of labour. As is well known, this was one of the main contributions of Adam Smith's *Wealth of Nations*, published in 1776. The resurgence of interest in economic growth motivates this article, which puts in an historical perspective the modern approach to the division of labour and growth, represented by Romer (1987). In particular we evaluate the modern approach from the perspective of Allyn Young's contribution, who developed the classical theory of Smith.

We are aware that, from a strictly methodological point of view, it may not be completely legitimate to compare a 'classical' theory to a recent formal model. Indeed, the former was expounded only verbally, in particular without the availability of modern mathematical tools. However, we attempt to identify the main characteristics of both approaches, which we argue correspond to quite different views of growth dynamics. In this respect, we assert that the modern theory does not completely capture the insights found in the classical theory, and that therefore the latter may still contribute to the development of a modern theory of economic growth.

A companion paper (Lavezzi, 2002) reconstructs the Smith-Young (and Marshall) theory showing that its main features are:

1. economic growth is endogenous: competition among profit-seeking entrepreneurs pushes them to continually reorganize their productive activities. This happens through two forms of division of labour: within and among firms. The former is characterized by the introduction of

organizational and technological changes, the latter by a change in the structure of the economy. These processes generate an endogenous increase in the network of interdependent economic units, allowing for the exploitation of productivity gains due to specialization. Economic growth is the aggregate result. Capital accumulation fuels this process and is inextricable from technological progress and population growth.

2. Growth has the nature of a cumulative, path-dependent process. Current opportunities for increasing productivity in the economy are determined by the degree of division of labour attained in the past.
3. Economic growth has the characteristics of a disequilibrium process, in which the data of the system, e.g. the number of commodities and the technologies available, are continuously changing under the pressure of competition. The economy therefore appears as an inherently unstable system, since qualitative changes constantly take place, activating productive and technological feedbacks. In this context, increasing returns are to be understood as 'generalized' or 'macroeconomic', being related to the size of the network of specialized, interdependent units. The growth of this network characterizes economic growth.

The latter two aspects make the classical approach essentially different from the modern one. Also, we maintain that the latter differs from the former in the following aspects: (*i*) in the bias towards the supply side and the partial neglect of demand; (*ii*) in an excessively important role attached to fixed costs; (*iii*) in the vague use of the term 'external economies', which may be instead replaced by 'network externalities'.

The rest of the paper is organized as follows: Section 14.2 introduces the Romer (1987) model; Section 14.3 critically compares the old and new approaches; Section 14.4 contains some concluding remarks.

14.2. GROWTH AND SPECIALIZATION IN THE EGT: THE ROMER MODEL

In this section we analyse the model by Romer (1987), which represents the formalization of growth and division of labour in the EGT. This model explicitly refers to Allyn Young (see Romer 1989, p. 108), but a reference to Young cannot exclude Smith. For this reason we consider the Romer model especially in the light of Young's contribution, but Smith is not overlooked.

Let us consider first the historical roots of the model, as presented by Romer (1989 and 1991). In particular, Romer (1991) outlines a history of growth theory, from Adam Smith to the EGT. He claims that in Smith two conflicting ideas coexist: the first is that competition ensures an efficient

allocation of given resources; the second is that growth depends on an endogenous process of accumulation.

In particular, competition conflicts with endogenous growth in the framework of the neoclassical model of Solow (1956). In the Solow model, output is produced by means of capital and labour, under constant returns to scale. In competitive markets factors receive in equilibrium the value of their marginal product and, by Euler's theorem, the product is exhausted by the remuneration of the factors of production. Hence, nothing can compensate for technological progress, the engine of long-run growth, which must necessarily be exogenous. Following Romer, since economists first developed the economic models of perfect competition for their mathematical simplicity, they were precluded from studying growth as an endogenous process.

The next step in Romer's reconstruction is Marshall's definition of external economies: they derive from the aggregate production of a commodity, and accrue to individual firms which remain price takers. Romer (1991, p. 87) remarks that these economies do not affect much of the neoclassical apparatus, but contribute to a better understanding of growth.

Then Romer (1991, p. 88) introduces Young's contribution, whose roots 'go back to Marshall and even Smith'. He argues (*ibidem*) that Young: '[l]ike Marshall, ..., called the beneficial effects arising from the introduction of a new good a positive external effect. Consequently, he tried to describe a model of growth driven by aggregate increasing returns that were external to individual firms'. Also, according to Romer (1989, p. 108) 'Marshall and Young choose to describe specialization in terms of competitive equilibrium with externalities'. Romer specifies that the introduction of new goods is not strictly equivalent to a Marshallian external economy (like 'trade knowledge'), but its consideration can lead to models which behave exactly like models with true externalities. Therefore, when the focus is on the introduction of new goods, fixed costs enter the picture because it is reasonable to assume the presence of a fixed cost when a new production is started. For Romer (1989, p. 108), Marshall and Young's story would be told in a 'more rigorous way in a model with fixed costs'. Such fixed costs may be interpreted as the costs necessary for the introduction of the commodity (e.g. research costs).[1] Once incurred, they do not contribute to the marginal cost of producing the good. Since in a competitive equilibrium the firm equates marginal cost to price, and therefore would be unable to recover the fixed cost, what is needed is a departure from the competitive framework.

Romer resorts to monopolistic competition, where (new) goods are differentiated but competition exists from potential producers, firms have market power but earn zero profits in equilibrium. Once this is admitted, the delay in the exploitation of Smith and Young's ideas is explained by the technical difficulties involved in building dynamic, general equilibrium macro-models with non-competitive sectors.

Therefore, it seems that the main obstacles for the exploitation of certain 'old' insights relate to mathematical difficulties. In particular, they depend on the requirement to represent growth dynamics as a general equilibrium process.[2] However, we argue that the interpretation of Smith and, especially, Young as advocates of the equilibrium approach may be challenged.

Let us briefly present the building blocks of the Romer (1987) model. The economy has two sectors: a final good sector and an intermediate goods sector. Only the final good is consumed. Intermediate goods are produced with the same technology using a capital good Z, owned by consumers in a given quantity. Production of intermediate goods entails a quasi-fixed cost, that is, no production at zero costs is feasible. The final good is produced under constant returns to scale, using intermediate goods and labour. In the intermediate sector a regime of monopolistic competition prevails: firms have market power but earn zero profits in equilibrium.

What is relevant is the functional form describing final good production, which must be such that 'having more available [intermediate] goods is useful' (Romer, 1989, p. 108). This can be obtained with the following production function:[3]

$$Y = L^{1-\alpha}\int_{\Re_+} x(i)^{\alpha}\, di \tag{1}$$

Here Y is the final good, L is labour, $x(i)$ is the quantity of the good i, and $0<\alpha<1$. Thus the marginal product of each intermediate good is decreasing.[4]

If all goods are produced in the same quantity $\bar{x} = N/M$ (which is the case here because of the symmetry of the model), where N is the total amount of intermediate goods and M is the range of goods actually produced, the production function becomes:

$$Y = L^{1-\alpha} N^{\alpha} M^{1-\alpha} \tag{2}$$

In equation (2) output can increase without bound with M, given N and L.

Aggregate production appears as if increasing returns were present, which is not the case if one considers equation (1). The range of intermediate inputs is theoretically infinite but the assumption of a fixed cost in terms of Z, whose quantity is given, guarantees that it is finite. The integration of a power function in equation (1) is the specific form in which intermediate goods are assembled in this model for production of the final good.[5]

To generate dynamics in this model, it is necessary to establish a mechanism which supports a growing M. This is obtained by assuming that Z is accumulated from forgone consumption by a representative individual who maximizes a discounted flow of utility.[6] That is, consumers' savings are invested in Z and rented to the intermediate goods sector. Consumers also inelastically supply a fixed amount of labour to the final good sector.

Romer specifies a particular form of the function $g(.)$ and of the cost function for the intermediate goods producers.[7] Then he shows that: (*i*) the quantity of the intermediate goods produced in equilibrium is $\bar{x} = 1$; (*ii*) the equilibrium condition for the monopolistically competitive sector is: $M(t)=Z(t)$; (*iii*) the following equation:

$$\frac{\dot{Y}}{Y} = \frac{\dot{c}}{c} = \frac{\dot{Z}}{Z} = \frac{1-\rho}{\sigma} \tag{3}$$

represents the solution for the consumer problem. Here σ is the reciprocal of the elasticity of intertemporal substitution and ρ is the intertemporal discount rate.

When $\sigma = 1$, Romer obtains that the consumption level in equilibrium is $c(t)=(G+\rho)Z(t)$ so that an increase in impatience leads to an increase in the level of consumption, a decrease in the level of savings and a reduction of the long-run growth rate.[8] Finally, Romer (1987, pp. 61–2) points out that 'this model is not one with a true positive externality, but it nonetheless behaves exactly as if one were present ... the economy will behave as if there is a form of exogenous, labour augmenting technological change'. He shows that equation (2) can be rewritten as:

$$Y(t) = M(t)^{1-\alpha}(L(t)^{1-\alpha}N(t)^{\alpha}) = AZ(t)\ L(t)^{1-\alpha} \tag{4}$$

where the constant A collects all the other constants.

In equation (4) the production function for aggregate output, although postulated as a constant returns to scale function, actually appears as if an external effect were present. Normalizing L to 1 returns the form of the familiar AK function which can be considered as the base for: 'the simplest endogenous growth model' (Barro and Sala-i-Martin, 1995, p. 38).

The next section analyses the elements which inspired Romer's formal model, along with a critical assessment of its relation with the Smith–Young theory of the division of labour and growth.

14.3. EVALUATION OF THE ROMER MODEL

In this section we discuss the Romer model in the perspective of the Smith–Young theory. We argue that the Romer model is a successful attempt to bring some relevant insights of such theory in the framework of EGT models, but at the same time it differs from the ‘old’ perspective in important respects.

It is fair to say that Romer himself is often cautious as to his simplifying hypotheses, but it seems that some of his claims cannot be safely taken for granted, in particular when he refers to Allyn Young. Romer follows Young with respect to a possible formalization of the way in which an increase in the ‘roundaboutness’ in production can increase the growth rate. This produces a sort of ‘macroeconomic’ increasing returns, according to the hypothesis on the imperfect substitutability of intermediate goods. However, it appears that this is done in a different perspective from Young’s.

In particular we mostly concentrate on a comparison between Romer and Young, it being understood that the latter’s contribution is related to Smith’s. We organize the discussion around four points:

1) Romer chooses an equilibrium approach against the disequilibrium approach of Young (and Smith). This is connected to the view of growth as a path-dependent process.
2) The Romer model is essentially supply-oriented and demand does not play an essential role as in Young (and Smith).
3) The emphasis on fixed costs is different.
4) Young was more cautious than Romer on the use of the concept of Marshallian external economies.

14.3.1. Equilibrium or Disequilibrium?

First of all, the Romer model is cast into an intertemporal equilibrium setting, while a reading of Young (1928) highlights that he strongly rejected the equilibrium approach to study endogenous economic growth. Young seemed on the contrary to advocate a disequilibrium theory of endogenous growth.[9]

Young (Ibidem, p. 528), discusses the Marshallian dichotomy on internal and external economies, arguing that the economies of a firm depend on what happens in an 'obscurer field' where: 'new products are appearing, firms are assuming new tasks, and new industries are coming into being'. Hence

> *No analysis of the forces making for economic equilibrium, forces which we might say are tangential at any moment of time, will serve to illuminate this field, for movements away from equilibrium, departures from previous trends are characteristics of it.* Not much is to be gained by probing into it to see how increasing returns show themselves in the costs of individual firms and in the prices at which they offer their products (italics added).

This view on disequilibrium is related to the very specific idea of Young on the functioning of a market economy. First, for Young (Ibidem, p. 531) the

economies: ‘which manifest themselves in increasing returns are the economies of capitalistic or roundabout methods of production’. Young focuses on one aspect of the general process of the division of labour, that is, the introduction of specialized machinery when labour has reached a certain degree of simplification, and the economies which stem from this process.[10] In particular, such economies depend on ‘large production’ and not on ‘large-scale production’. Attention should therefore be placed neither on individual firms, and on their negatively sloped cost curves, nor on individual industries. Young introduces what may be termed ‘macroeconomic increasing returns’,[11] which can be understood from an analysis of the entire economy, considered as a large interactive system. The market is in fact defined by Young (Ibidem, p. 533) as: ‘an aggregate of productive activities, tied together by trade’.

Young (Ibidem, p. 531 and 533) takes an ‘inclusive view [of the market, which is not] an outlet for the products of a particular industry, and therefore external to [an] industry, but [i]s the outlet for goods in general. [Therefore:] the size of the market is determined and defined by the volume of production’. This immediately leads him to this reformulation of Smith’s theory: the division of labour is limited by the division of labour. Although reminiscent of Say’s law, this argument is more far-reaching:[12] it asserts that both demand and supply are endogenously determined according to the degree of division of labour prevailing.[13] This amounts to recognizing that the extent of the market is at least partially endogenous.

So, we arrive at the important implication that

> the counter forces which are continually defeating the forces which make for economic equilibrium are more pervasive and more deeply rooted in the constitution of the modern economic system than we commonly realise. Not only new or adventitious elements, coming in from the outside, but elements which are *permanent characteristics of the ways in which goods are produced* make continuously for change. Every important advance in the organisation of production ... alters the conditions of industrial activity and initiates responses elsewhere in the industrial structure which in turn have a further unsettling effect. This change becomes progressive and propagates itself in a cumulative way (Ibidem, p. 531. Italics added).

Thus for Young, not only is economic growth endogenous, but also endogenous forces generate disequilibrium, in the sense that, in the growth process, the structure of the economy and the technological opportunities cannot *a priori* be considered fixed. Young (Ibidem, p. 533) the apparatus of supply and demand is incapable of exploring this sort of dynamics, since it may ‘divert attention to incidental or partial aspects of a process which ought to be seen as a whole’.[14] The use of an equilibrium approach to study growth seems even to imply the impossibility of defining growth as an endogenous process (Ibidem, p. 535).

Young (Ibidem, p. 534) then introduces the concept of reciprocal demand. Reciprocal demands among firms are characterized by a certain level of elasticity, to be interpreted as the capacity for the increased production of a good to generate demand for other goods.

The elasticities are different for different products, so growth in the economy will be different among sectors. In any case '[e]ven with a stationary population and in the absence of new discoveries in pure or applied science there are no limits to the process of expansion except the limits beyond which demand is not elastic and returns do not increase' (Ibidem, p. 534).

Differently, in the representation of the productive process, Romer adopts an equilibrium approach, preserving the 'one-way avenue' from given resources to final output, although by means of an intermediate sector. It is not clear that this can be taken as a faithful representation of the economy which Young had in mind. He considered the economy as an 'interrelated whole', where the extent of the market is endogenous and feedbacks, for instance in the form of 'reciprocal demands', among productive units are continuously displacing the tendency towards equilibrium, when this is interpreted in an allocative sense.[15]

In the Romer model, there are two allocation problems: the first regards the allocation of the given resource Z among the intermediate goods producers; the second is the allocation problem of consumers between consumption and saving. The first problem is solved imposing the zero profit condition in the intermediate sector: in this case the results is an equilibrium range of intermediate goods. The second is solved by utility maximization of consumers, given the paths of the rental price for Z and the price for the consumption good.

In any case, production is never assumed to take place under increasing returns, due for instance to the continuous re-organization of the production process, to learning by doing, to improvements in the technology, as emphasized by Smith and Young (see Lavezzi, 2002). In the Romer model increasing returns appear in the aggregate, as for Young, but they are generated by a series of equilibrium conditions and depend on a particular hypothesis on the way intermediate inputs are assembled.

In the Introduction we also argued that in the Smith–Young framework growth based on the division of labour is a path-dependent, cumulative process. This nature of growth emerges from the various quotations from Young, and can be inferred even from Adam Smith. In Smith it is capital accumulation that transforms economic growth based on the division of labour into a *cumulative process*: the division of labour is allowed by the

accumulation of capital, and cannot proceed unless some previous stage of division of labour has been reached. Therefore, growth is characterized by path dependency: current opportunities to increase productivity by further subdivisions of labour, depend on the past stages of the division of labour.[16]

Is there path dependency in the Romer model? The answer is in the affirmative in the sense that the model predicts lack of convergence, since economies starting poor because they do not specialize stay poorer than economies that start richer, because they produce a larger range of intermediate goods. However, this type of model does not feature path dependency in the strict sense, since cumulative effects are absent: if the saving rate, e.g., increases through a decrease in an exogenous factor like the intertemporal discount rate, the range of intermediate goods increases as well as the growth rate, but this *per se* does not preclude the saving rate, and the range of intermediate goods, from decreasing in a subsequent period due to a change in the same parameter of the opposite sign.

14.3.2. Supply-side or Demand-side?

In the Romer model consumers save and invest in Z; this increases the range of intermediate goods which in turn increases production and income. In the intermediate sector there are firms potentially active, but the decision of such firms to produce is not due to a sudden increase in demand for their good. *A potential demand is always existing for an infinite number of intermediate goods*, because of the form of the production function for final goods. Potentially active firms can become operative once the available quantity of Z makes it possible; thus it is savings that foster growth.

The increase in specialization, i.e. in the number of intermediate goods, which is permitted by savings, increases Y, which is income earned by consumers and subsequently consumed or invested. Thus, the growth of Y is constrained by the supply of intermediate goods, in turn constrained by the availability of the primary resource Z.

Let us briefly recall the role of demand in the process of growth based on the division of labour. It is well known that in Smith and Young the main limit to the division of labour, and thus to growth, is the extent of the market.

According to Smith (WN I.i.1), a 'natural' predisposition for socio-economic interactions allows individuals to specialize and obtain a gain from trading their surplus products, i.e. the production in excess of their own consumption, deriving from higher productivity due to the specialization. The extent of the market comes into play here: an individual has the incentive to specialize if he possesses 'power of exchanging' that surplus, i.e. if sufficient demand exists, allowing the agent to purchase other goods with the revenue from the disposal of his surplus product.[17] Hence, economic growth may be spurred by the creation of a network even of similar (i.e. not specialized) individuals. Once connected, they will sort themselves out in different occupations, and increase their aggregate production. Thus, in the Smithian framework, demand and supply are at least on the same level.[18]

Can we find evidence in the Romer model for the principle that 'the division of labour is limited by the extent of the market'? What is certainly true is that the division of labour, that is the introduction of new intermediate goods in the production of the final good, is limited by fixed costs. However, when discussing the introduction of new, specialized machines (which corresponds to the intermediate goods in the Romer framework) as an aspect of the process of the division of labour, Young mainly placed emphasis on another question: that is on the possibility of adopting more capital-intensive, highly productive methods, conditional on the possibility of selling a large output In this case the absence of demand limits the division of labour; Young seemed to be less concerned with resource-constraints faced by firms.

14.3.3. What Role Do Fixed Costs Have?

The latter point is also related to the question of the role of fixed costs in Young's theory. Young seemed to be aware of them in the discussion of the introduction of machines, but did not emphasize their role. Sandilands (2000, p. 315) comments: 'Young did not say that specialization is limited by the presence of fixed costs, though he did say that specialization increasingly took the form of greater roundaboutness in the economy as whole. In his theory, fixed costs and increased roundaboutness are not so much a constraint on growth as its consequence'.

The issue for Young (and Smith) was the creation of new markets or the extension of existing ones. Young (*ibidem*, p. 530) writes in a famous passage that 'it would be wasteful to make a hammer to drive a single nail' or 'to furnish a factory with an elaborate equipment of specially constructed [machines]' if the market is not sufficiently large. According to this idea, noted also by Kaldor (1972, p. 1242), the capital–labour ratio chosen by firms depends on the extent of the market and not only on relative factor prices. Hence, the importance of the extent of the market for the division of labour implies a role for the extent of the market in the choice of the optimal capital–labour ratio.

At any rate, in the context of the whole of Young's theory, his discussion on the introduction of new specialized machines seems more relevant for its 'dynamic' implications. Namely, for the choice of the firm's internal organization and technology, and for the feedbacks on the environment, for instance for capital–goods industries.

14.3.4. External Economies or Network Externalities?

Romer claims that, as reported in Section 14.2, Young (and Marshall) discussed growth and specialization in terms of competitive equilibrium with positive external effects. However, this interpretation does not seem to be correct. From the previous discussion, we observe first that Young firmly rejected the approach to the study of economic growth based on equilibrium of supply and demand.

At the same time he (1928, p. 528) considered the Marshallian distinction between internal and external economies, as implying just 'a partial view' on the growth process. For Young (1928, pp. 527–8) the Marshallian distinction is 'fruitful', but partially misleading.

It may well be that, when considering the positive effect generated outside a productive unit, in that 'obscurer field' as Young called it, a more appropriate concept than external economy for Young's theory is that of *network externality*. Consider the following definitions from Economides' *Dictionary of Terms in Network Economics*:[19] '*Networks*: networks are composed of complementary nodes and links ... *Network externality*: a network exhibits network externalities when the value of a subscription to the network is higher when the network has more subscribers'.

In the Smith–Young framework the creation of a network of specialized production units (i.e. 'complementary links'), or the increase in the number of 'subscribers' to an existing one, increases the 'value of the subscription' of the individual participant, in the sense that it may represent an increase of its market. The discussion by Young on reciprocal demands implies that the network itself is subject to endogenous change, as growth of production may lead to creation of new specialized sectors, activating feedbacks elsewhere, etc. The aggregate results should be endogenous economic growth.[20]

14.4. CONCLUDING REMARKS

Summing up the arguments put forward in the previous section, we believe that the Romer model did successfully capture some features of the growth dynamics generated by the division of labour. However, our re-evaluation of the contributions of Allyn Young and Adam Smith indicates that some aspects are in need of different formalizations.[21]

Both the old and new approach recognize the endogenous nature of economic growth. Both relate growth to an increasing complexity of the economy, and introduce a sort of 'macroeconomic' increasing returns. The main differences relate to the view of growth based on division of labour as an equilibrium or path-dependent, disequilibrium process. Other differences regard the role of demand, of fixed costs and the suitable notion of external effect. We therefore conclude that the classical theory of division of labour and growth may still provide insights for the development of a modern theory of economic growth.

NOTES

1. See also Romer (1990).
2. As Romer (1989, p. 70) writes: '[g]rowth is a general equilibrium process'.
3. Dixit and Stiglitz (1977) introduced this form in a utility function to express preference for variety in consumption. Ethier (1982) proposed to utilize it for a production function.
4. The more general form of equation (1) is:

$$\mathrm{Y} = \mathrm{L}\int_{\Re_+} g\left(\frac{x(i)}{L}\right)di$$

where it is required that $g(.)$ is increasing and strictly concave, with $g(0) = 0$.

5. Another functional form to aggregate intermediate goods appears in the literature, the CES specification. In the CES specification, the production function can be expressed by:

$$Y = L^{1-\alpha}\left\{\left[\int x(i)^{\theta}\,di\right]^{\frac{1}{\theta}}\right\}^{\alpha}$$

where θ is a parameter reflecting the elasticity of substitution among different intermediate inputs, given by $\varepsilon = 1/(1-\theta)$. When $0 < \theta < 1$, goods are imperfect substitutes (i.e. $1 < \varepsilon < \infty$); when $\theta = 1$ goods are instead perfect substitutes. Note that the formulation in equation (1) is simply obtained by putting $\theta = \alpha$. In the CES case, when all intermediate goods are produced in the same quantity $\overline{x} = N/M$, we obtain:

$$Y = L^{1-\alpha}N^{\alpha}M^{\alpha\cdot(1-\theta)/\theta}.$$

Thus output can increase without bounds in M as long as $0 < \theta < 1$. A general form for this type of production function is the following:

$$Y = L^{1-\alpha}\left[\int x(i)^{\theta}\,di\right]^{\gamma}\left[\int x(i)^{\delta}\,di\right]^{\tau}$$

This formulation preserves the homogeneity of degree α in $x(i)$ and the positive relation between M and Y, if the following conditions are satisfied:

i) $\gamma\theta + \tau\delta = \alpha$, $0 < \alpha$, $\theta < 1$;
ii) $\gamma + \tau > \alpha$.

The form chosen by Romer corresponds to the case in which: $\theta = \alpha$, $\gamma = 1$, $\tau = 0$. The case of a CES specification is obtained when: $\gamma = \alpha/\theta$, $\tau = 0$.

6. A conventional isoelastic utility function $U(c)$ is considered.
7. In particular, $g(.)$ is strictly concave on the interval $[0, x_0]$, and has a constant slope equal to 1 on the interval $[x_0, \infty]$. In addition: $g(0) = 0$ and $g'(x_0) = 1$. The intercept on the vertical axis obtained by prolonging the slope equal to 1, is indicated by G. The cost function is $h(x) = (1 + x^2)/2$.

8. This equilibrium growth rate is suboptimal: a policy which raises savings would positively affect long-run growth.
9. In Lavezzi (2002) we argue that also the original theory of Adam Smith can be interpreted in this way. See also Richardson (1975).
10. Young also treats the division of labour among industries.
11. This definition appears in Currie (1997) and is compatible with the one recently proposed by Buchanan and Yoon (2000, p. 45) with respect to Smith: 'the Smithean proposition that relates the division or specialization of labour to the extent of the market is best captured by the notion of *generalized increasing returns*, which implies only that the degree of specialization utilized increases with the size of the whole nexus of economic interaction thereby increasing the ratio of positively valued outputs to inputs' (Italics in the original text).
12. Young (1999b, p. 145) in fact criticises Say's law as such.
13. This aspect is assumed in recent models such as Yang (1999), where the agents are producer–consumers, and the structure of demand and supply is simultaneously determined with the degree of division of labour.
14. See also (Young, 1999a, p. 45): 'Seeking for equilibrium conditions under increasing returns is as good as looking for a mare's nest. Certainly the matter cannot be explained by this curve apparatus, which does not see things "in their togetherness"'
15. Moreover, interdependence among sectors in the Romer model appears in the sense that the final good is produced by means of intermediate goods in one period, and becomes a factor of production for them in the following period if not consumed. Again, this does not seem to be the story told by Young on the reciprocal effects triggered by increases in supply, which stimulate increases in demand, which in turn become increases in supply by other firms, etc.
16. Kaldor (1972, p. 1245, and 1975, p. 355) draws strong implications from the recognition of this aspect, denying the possibility of defining growth as an equilibrium process.
17. Clearly, the same logic applies to a firm: specialization of operations is profitable if there is demand for the higher quantity of goods that the firm can produce by specializing.
18. For further discussion of demand in the process of the division of labour see Lavezzi (2002).
19. Available at: http://www.stern.nyu.edu/networks/dictionary.html.
20. In the present case, the network effects should be understood in a particular sense, with respect to the existing literature. In fact, we know that the choice to specialize in the production of a particular good depends on others' choice to specialize in different goods. Thus we can call the benefit from specialization *conformity effects* (in the sense that different agents make the choice of specialization), keeping in mind that the specialization is in different activities. Typically, conformity effects are such that the benefits of an action, for example the adoption of a technological standard (see, e.g., Arthur, 1987), increase in the number of those making the same choice. In this sense the *positive feedback effect* (see Agliardi, 1998) reinforcing the outcome of agents' choices should be understood correspondingly: the choice of agents to specialize is mutually reinforcing, but they are involved in different activities, and therefore, e.g., adopt different technologies.
21. For instance, the comment provided by Heal (1999, p. xxiii) supports this claim. Heal, after presenting the main features of Young's growth theory, writes: '[t]his seems an interesting intuition, broadly consistent with casual empiricism, and *not captured by any formal growth models*. It has some resemblance to evolutionary models in biology, where evolution leads to increasing complexity and longer food chains' (Italics added). Moreover, from an historical point of view, the Romer model of growth and specialization has been recently criticized by Sandilands (2000, p. 315), for not being able to: 'fully capture Young's view of the links between fixed costs, specialization, external economies, and the economy-wide external returns that make growth a semi-automatic, self-perpetuating process'.

15. The interaction between growth and cycle in macrodynamic models of the economy

Serena Sordi

15.1. INTRODUCTION

Since the beginning of the study of the dynamics of industrial capitalism, economic growth and business cycles have been seen as indissolubly linked. In this regard, it is important to remember that Marx (1954, ch. 25) – 'the first economist seriously to study the cyclical aspect of capitalism' (Goodwin, 1986, p. 15) – considered the business cycle as nothing other than the basic way in which capitalist economies develop, due to the interaction between the accumulation process (and the resulting growth of productive capacity) and the conflict over income distribution between capitalists and workers. Schumpeter (e.g., 1939), in turn, produced an integrated theory of growth and business cycles in which economic fluctuations are nothing other than the 'form which progress takes in capitalist society' (Schumpeter, 1927, p. 295).

As is well known, a crucial role in Marx's (endogenous) explanation of the mutual conditioning of growth and cycle is played by the size of the *reserve army of labour*. In short, the story runs as follows (see Marx, 1954, p. 597). In periods of high rates of accumulation, the reserve army of labour shrinks, thus leading to an increase in labour's bargaining power. This, in turn, causes a change in income *distribution* in favour of workers, entailing a decline of profits and a consequent decline of accumulation. But this leads to an increase in unemployment and pushes the wage share down. As a consequence, the profitability of real investment is restored, the rate of capital accumulation goes up and the sequential mechanism just described can start again. The crucial role in Schumpeter's explanation, on the other hand, is played by the concept of the 'pioneering entrepreneur' who carries out an *innovation*, whether this is a new method of production or transportation, a change in industrial organisation, a new product, the opening up of a new market or a new source of materials. This opens up new

profitable avenues such that more entrepreneurs are induced to innovate, giving rise to an investment boom and driving growth for the economy as a whole. Once the innovation is fully exploited, however, the economy relapses into a depression which lasts until the accumulation of new ideas creates a favourable climate for a new burst of innovating investment and so on.

Both Marx's and Schumpeter's theories, however, are descriptive rather than analytical. Thus, the question of how to incorporate the Marxian 'distributional' mechanism, or the Schumpeterian 'innovational' mechanism or a combination of the two in a formalized dynamic model of the economy does not have a single, straightforward answer.[1]

Keeping this in mind, in the present paper we intend to review, with regard to their capacity to represent the cyclical growth of the economy, some early contributions to *non-market clearing* macrodynamics.[2] This will be done by analysing two prototype models. In Section 15.2, the simple linear Keynesian ('multiplier–accelerator') model is presented and analysed with respect to its cyclical growth properties. Two results easily follow from a qualitative analysis of this model: first, there cannot be any interaction between the cyclical and the growth component of the solution; second, even as a 'pure' cyclical model, it is not at all satisfactory given that the representation of persistent fluctuations can only be obtained for a very special combination of parameter values. Our purpose is to show that a non-linear version of the model, in addition to solving the second of the two problems, provides some interesting hints also with regard to the first, once we assign some role to 'innovational' investment in the Schumpeterian sense.

In Section 15.3 we outline Goodwin's growth cycle model which provides a formalisation of the Marxian reserve army mechanism described above. The original version of this model is known to suffer from two weaknesses. First, it is structurally unstable, in other words, even the smallest perturbation of its structure can destroy the closed orbit (cyclical) character of its solution; second, it neglects altogether any disequilibrium in the product market. To conclude this section, we briefly discuss a modified version of the model which, in relaxing Goodwin's extreme ('classical') assumption about savings, is shown to produce limit cycle solutions, thereby overcoming the first of the two weaknesses. In the attempt to overcome also the second, an integration of the two prototype models is then sketched in Section 15.4. Finally, Section 15.5 contains some conclusions and suggestions for further research.

15.2. THE KEYNESIAN ‘MULTIPLIER–ACCELERATOR’ MODEL

15.2.1. A Linear Formulation of the Model

In most of the earlier macrodynamic models which followed the publication of Keynes’ *General Theory* (1936), the prevalent attitude was that of a separate handling of business cycles and growth. These were typically linear models based on purely endogenous relationships, aimed at explaining the aggregate behaviour of consumers (through the *multiplier*) on the one hand and that of entrepreneurs (through some version of the *principle of the accelerator* or some other theory of aggregate investment) on the other. The dynamics resulting from such an interaction between the multiplier and the accelerator is either cyclical or monotonic and does not succeed therefore in replicating the observed persistent cyclical growth of real economies. The only way out of this puzzle is to assume that the parameters of the model are such that the solution is cyclical (with fluctuations of constant amplitude) and then to add to the model an autonomous component of aggregate expenditure, which grows exogenously through time. In this case, the ‘multiplier–accelerator’ interaction implies constant amplitude fluctuations of output around a growing trend, but, *by construction*, there cannot be any interaction between the growth and the cyclical components of the dynamics.[3]

All this can be easily illustrated by using the following (prototype) linear model of the ‘multiplier–accelerator’ interaction:[4]

$$C = cY,\ 0 < c < 1 \tag{1}$$

$$I = I_i + I_a(t) \tag{2}$$

$$\dot{Y} = \frac{1}{\varepsilon}\left[(C+I) - Y\right],\ \varepsilon > 0 \tag{3}$$

$$\dot{I}_i = \frac{1}{\theta}\left(\overline{v}\dot{Y} - I_i\right) = \frac{1}{\theta}\left\{\overline{v}\dot{Y} - \left[I - I_a(t)\right]\right\},\ \overline{v}, \theta > 0 \tag{4}$$

Equation (1) is a standard Keynesian consumption function, where C denotes aggregate consumption and Y the national income. Equation (2), on the other hand, simply states that total investment (I) is the sum of an induced component (I_i) and an autonomous component ($I_a(t)$). The basic dynamic ingredients of the model are the two *error-adjustment mechanisms* (3) and (4), according to which the aggregate supply and the induced investment adjust to their desired levels, determined by total demand ($C + I$)

and the accelerator principle ($\overline{v}\dot{Y}$) respectively, where $\overline{v}$ is the (constant) desired capital–output ratio. In writing (3) and (4), we have assumed that both error-adjustment mechanisms involve a simple exponential lag, of lengths equal to ε and θ respectively (see Allen, 1967, pp. 76–9 and 94–5).

From (3) and (4), given (1) and (2), we obtain the following second-order differential equation, the solution of which describes the dynamics of the national income

$$\ddot{Y} + \frac{1}{\varepsilon}(\varepsilon + s - \overline{v})\dot{Y} + \frac{s}{\varepsilon}Y = \frac{1}{\varepsilon}\left[\dot{I}_a(t) + I_a(t)\right] \tag{5}$$

where $s = 1 - c$ and where we have chosen the time unit so as to have $\theta = 1$.

In the case in which $I_a(t) = 0$ for all t, (5) is a homogeneous differential equation. By a simple analysis of the roots of its characteristic equation, we can then conclude that the relation between the values of the parameters of equation (5) and the dynamics of the model is the one summarized in Table 15.1 and shown graphically, for two different cases, in Figure 15.1.[5] As anticipated, the resulting dynamics is *either* monotonic (for combinations of the parameter values in regions ***A*** or ***E***) *or* oscillating (for combinations of the parameter values in regions ***B***, ***C***, or ***D***).

Table 15.1 – Intervals of parameter values and type of solution

Intervals of values of $\overline{v}$	*Type of solution*
$\overline{v} \le (\sqrt{\varepsilon} - \sqrt{s})^2$	monotonic, convergent (***A***)
$(\sqrt{\varepsilon} - \sqrt{s})^2 < \overline{v} < \varepsilon + s$	oscillating, damped (***B***)
$\overline{v} = \varepsilon + s$	oscillating, constant amplitude (***C***)
$\varepsilon + s < \overline{v} < (\sqrt{\varepsilon} + \sqrt{s})^2$	oscillating, divergent (***D***)
$\overline{v} \ge (\sqrt{\varepsilon} + \sqrt{s})^2$	monotonic, divergent (***E***)

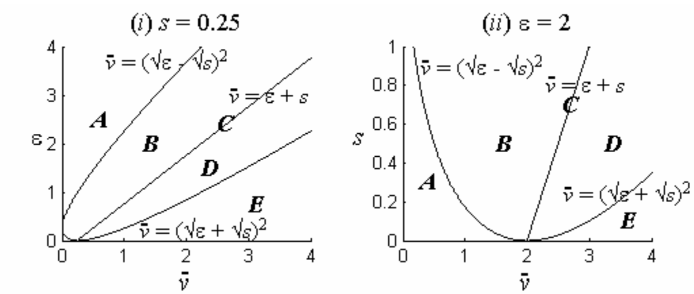


Figure 15.1: Regions of parameter values and type of solution

On the other hand, if we assume that there is an autonomous component of investment which increases over time, for example, such that

$$I_a(t) = a_0 + a_1 t,\ a_0,\ a_1 > 0$$

from (5) we obtain the following non-homogeneous differential equation

$$\ddot{Y} + \frac{1}{\varepsilon}(\varepsilon + s - \overline{v})\dot{Y} + \frac{s}{\varepsilon}Y = \frac{1}{\varepsilon}(a_2 + a_1 t) \tag{6}$$

where $a_2 = a_0 + a_1$. In this case, choosing a combination of values for the parameters of equation (6) along the straight line ***C*** of Figure 15.1, we obtain the representation of (constant amplitude) fluctuations around a growing trend.[6]

By construction, however, there cannot be any interaction between the cycle component (Figure 15.2(*i*)) and the growth component (Figure 15.2(*ii*)) of the dynamics in that they are simply superimposed (Figure 15.2(*iii*)).[7]

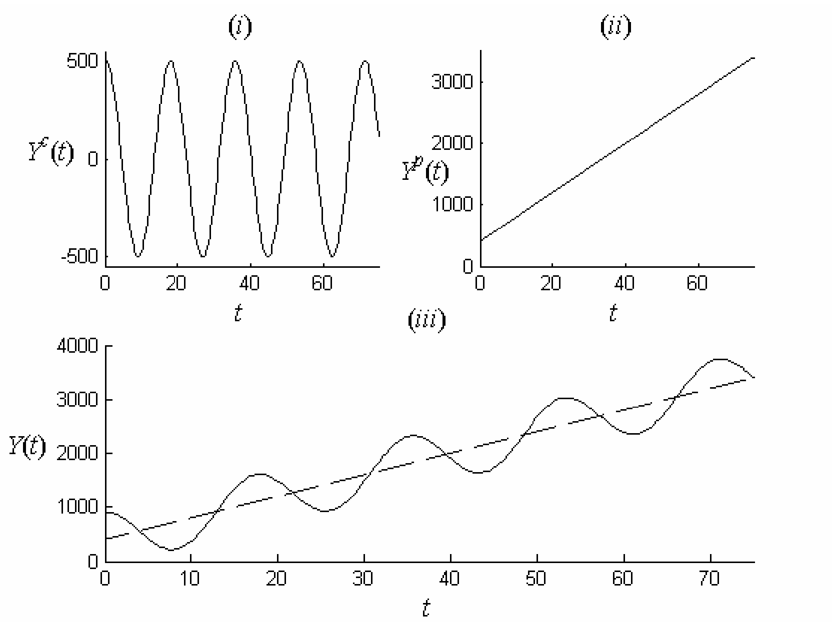


Figure 15.2: Dynamics of the linear multiplier–accelerator model with a linear trend and oscillations of constant amplitude (for $\varepsilon = 2$, $s = 0.25$, $\overline{v} = 2.25$, $a_0 = 90$, $a_1 = 10$)

15.2.2. A Non-linear Formulation of the Model with 'Innovational' Investment

The previous analysis demonstrates that the simple 'multiplier–accelerator' model cannot produce both growth and cycle: in the case considered, for example, it proves to be a pure cyclical model, such that the representation of cyclical growth can only be achieved by superimposing a trend on it.[8] On the basis of what was stressed in Section 15.1, however, this is hardly surprising given that in such a model (*i*) only the product market is formalised, whereas the labour market is neglected altogether and (*ii*) any kind of 'innovational' investment in the Schumpeterian sense is ignored. Given (*i*), it appears that the consideration of the conflict-over-distribution mechanism would certainly require a complete reformulation of the model and we will not attempt such a task here. Thus, it remains to be seen whether it is possible to incorporate into the model the other – 'Schumpeterian' – mechanism.

In order to do that, let us consider a non-linear version of the model (see Goodwin, 1951), which at least enables us to represent persistent fluctuations (limit cycles) for a larger range of parameter values. This is easily done by replacing the desired level of investment $\overline{v}\dot{Y}$ in the error-adjustment mechanism (4) with the non-linear function $\phi(\dot{Y})$, where ϕ is as shown in Figure 15.3(*i*). The rationale for the 'sigmoid' shape is the fact that investment in fixed capital has an upper limit ($I_i^{\max}$), given by the maximum rate of investment allowed by existing productive capacity, and a lower limit ($-I_i^{\min}$), corresponding to zero gross investment. Thus, it is plausible to assume that the acceleration principle determines the desired level of the capital stock only over some middle range, but passes to complete inflexibility at either extreme.[9] Equation (5) becomes

$$\ddot{Y} + \frac{1}{\varepsilon}\left[(\varepsilon + s)\dot{Y} - \phi(\dot{Y})\right] + \frac{s}{\varepsilon}Y = \frac{1}{\varepsilon}\left[\dot{I}_a(t) + I_a(t)\right] \tag{7}$$

From (7), assuming (as in Goodwin, 1951, p. 12) that autonomous investment is constant and equal to I_a^* for all *t*, we obtain

$$\ddot{Y}^c + \frac{1}{\varepsilon}\left[(\varepsilon + s)\dot{Y}^c - \phi(\dot{Y}^c)\right] + \frac{s}{\varepsilon}Y^c = 0 \tag{8}$$

where $Y^c = Y - I_a^*/s$. As is well known (see, for example, Gandolfo, 1997, pp. 440–1), in the case in which $\overline{v} > \varepsilon + s$ (i.e., the equilibrium is locally unstable) the solution of equation (8) is a *limit cycle*, describing *persistent* fluctuations of national income around the (constant) equilibrium I_a^*/s (see Figure 15.4).

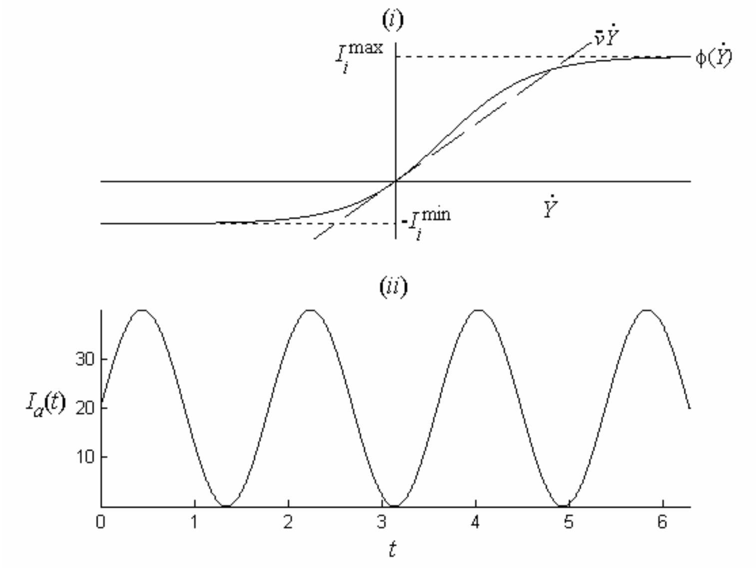


Figure 15.3: (i) The induced and (ii) the autonomous component of investment

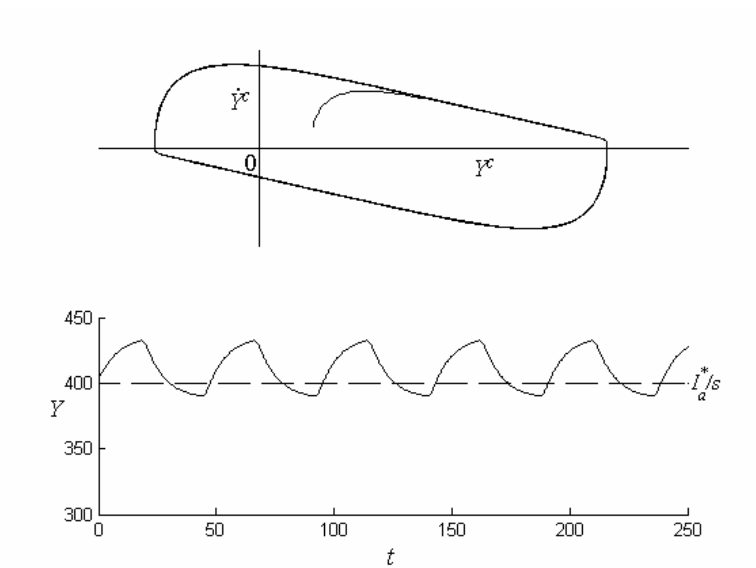


Figure 15.4: The limit cycle of the non-linear accelerator model (for $\varepsilon = 2$, s = 0.25, $\bar{v} = 3$, $I_a^* = 100$)

To assess the importance of this result, however, it is useful to go back to the meaning of *autonomous* – as opposite to *induced* – investment. In the model we are considering, such a component includes all that investment in fixed capital that is not explained by the acceleration principle, mainly, therefore, *innovational* investment in the Schumpeterian sense. Thus, it is not at all satisfactory to assume that it is constant over time. Rather, a better, although rough, way of formalizing Schumpeter's idea of clustering of innovations is to assume, as suggested by Goodwin (1946, p. 97), that such a component of investment is a *periodic* function of time of the kind

$$I_a(t) = b\left[1 + \sin\left(ct\right)\right] \tag{9}$$

where the two parameters *b* and *c* determine the amplitude and the frequency of innovational investment respectively (see Figure 15.3(*ii*)).

An example of the simulation of the model with the 'forcing' effect of innovational investment (9) is given in Figure 15.5. As we see, the result is now that the autonomous ('innovational') investment interacts with the cyclical dynamics of the model in such a way as to generate cyclical growth, although, for the parameter values chosen, the latter is only *transient*.[10]

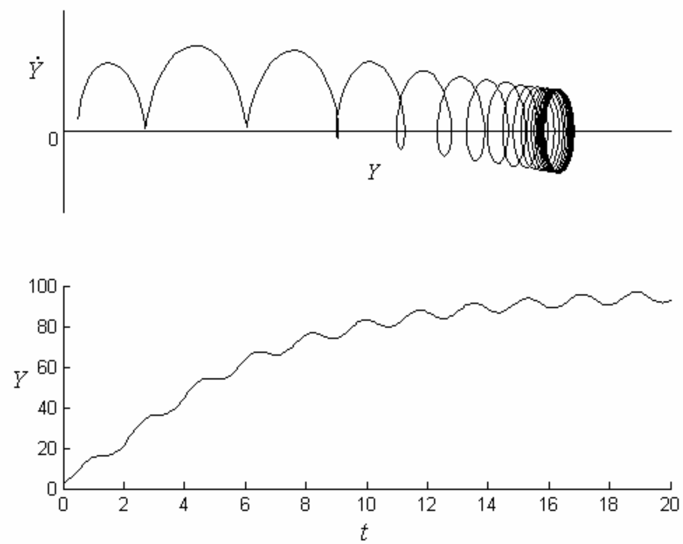


Figure 15.5: Dynamics of the non-linear accelerator model with innovational ('forcing') investment (for $\varepsilon = 2$, $s = 0.25$, $\bar{v} = 3$, $b = 20$, $c = 3.5$)

15.3. THE 'MARXIAN' GROWTH CYCLE MODEL

Since the late Sixties, a different approach to growth cycles has been developed, based on Goodwin's (1967) growth cycle model.[11] We call this approach 'Marxian', in order to stress that in it the crucial role is played by the Marxian reserve-army-of-labour mechanism we described in the introduction. Taking account of the vast literature that has appeared since the publication of the original contribution (hereafter, OVM = Original Version of the Model), in what follows we propose and analyse two generalisations of the OVM.

Before doing that, let us briefly recall that the OVM gives rise to the following dynamical system of the Lotka–Volterra type

$$\dot{E} = (g - g_n)E = \left(\frac{1}{\overline{v}} - g_n - \frac{1}{\overline{v}}U\right)E \qquad (10)$$

$$\dot{U} = \left[f(E) - \alpha\right]U \approx \left[-(\gamma + \alpha) + \rho E\right]U \qquad (11)$$

where, apart from the notation already introduced in the previous section, L stands for employment, $A = Y/L$, labour productivity, N, the labour force, $g_n = \alpha + \beta$, the natural rate of growth, W, the real wage, $U = WL/Y = W/A$, the share of wages, $E = L/N$, the employment rate, S, total savings, and g, the rate of growth of output, and where it is assumed that

$$A = A_0 \exp(\alpha t)$$

$$N = N_0 \exp(\beta t)$$

Equations (10) and (11) easily follow from these basic assumptions together with the assumption of a Phillips curve for the *real* wage dynamics

$$\hat{W} = f(E) \approx -\gamma + \rho E,\ \gamma, \rho > 0$$

and a classical assumption about savings behaviour, according to which all profits are saved and invested and all wages consumed.

As is well known, the solutions in E and U of equations (10)–(11) are cyclical (the positive equilibrium point being a *centre*) (see Figures 15.6(*i*) and 15.6(*ii*)) and, given that $g = (1-U)/\overline{v}$, this implies that the rate of growth of the economy is also cyclical (Figure 15.6(*ii*)). Thus, output is subject to growth cycles as shown in Figure 15.6(*iii*).

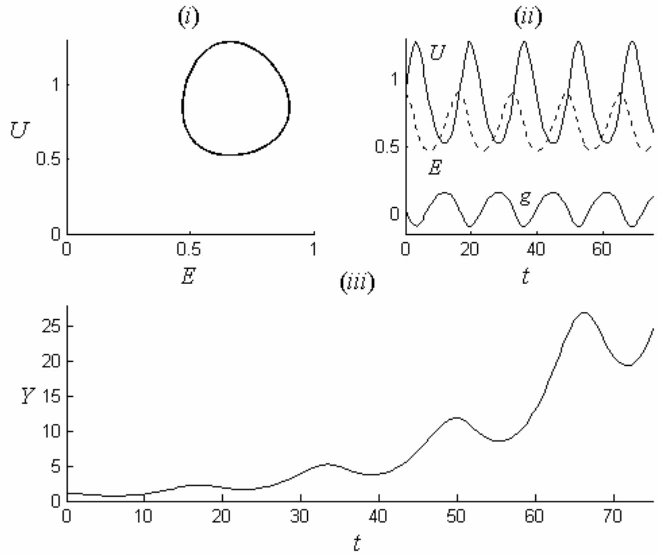


Figure 15.6: Growth cycles of the OVM

It is obvious that, in this case, the capacity to generate growth cycles is intrinsic to the model and is not due to the choice of any particular form for the functions of the model. We believe, however, that two aspects reduce the importance of this result. First, the fact that the positive equilibrium point of the model is a centre implies that the fluctuations of E and U around their equilibrium values are of an amplitude which fully depends on initial conditions. In the existing literature on the topic, however, one can find a number of contributions in which the OVM is modified in such a way as to generate limit cycle dynamics of the relevant variables. We present and discuss one of these extensions of the OVM in Section 15.3.1. Second, and more importantly, the OVM neglects altogether any effective-demand considerations. There cannot be, therefore, any role, in the generation of the growth cycles, for adjustments to product market disequilibrium. This is in sharp contrast with the Keynesian model presented in Section 15.2. A 'hybrid' version of the OVM, which attempts to introduce a Keynesian flavour *while preserving the capacity of the model to generate growth cycles*, is outlined in Section 15.4.

15.3.1. A Modified Version of the Model with Differential Savings

In Sordi (2001), the author proposed a modified version of Goodwin's model in which the 'distributional' mechanism plays a role in the dynamics also *via* savings behaviour. She studied the case in which both capitalists and workers save a fraction of their incomes and showed that the model can produce persistent oscillations (limit cycles) and even chaotic dynamics if, in introducing such a modification of Goodwin's classical assumption, (*i*) we take account of Pasinetti's criticism (e.g. 1962) of Kaldor's (1956) approach to differential savings; and (*ii*) we consider a more general version of the Phillips curve, according to which the rate of growth of real wages depends not only on the *level* of the rate of employment, but also on its *rate of change*.

To introduce this first 'Modified Version of the Model' (MVM1), with the help of some additional notation, we write

$$S_c = s_c P_c = s_c r K_c$$

$$S_w = s_w\left(WL + P_w\right) = s_w\left(WL + rK_w\right)$$

$$Y = WL + P_w + P_c$$

$$r = \frac{P}{K} = \frac{Y - WL}{K} = \frac{1-U}{\overline{v}}$$

$$g = \frac{s_w + \Delta s\left(1-U\right)X}{\overline{v}}$$

and[12]

$$\hat{W} = f_1\left(E, \hat{E}\right) = h\left(E\right) + \delta\hat{E},\ h'\left(E\right) > 0,\ h''\left(E\right) > 0,\ \delta > 0 \tag{12}$$

where $X = K_c/K$ is the proportion of capital held by capitalists, P_w, workers' profits, P_c, capitalists' profits, $P = P_c + P_w$, total profits, $r = P/K$, the rate of profit, s_w and S_w, workers' propensity to save and savings, s_c and S_c, capitalists' propensity to save and savings, $S = S_w + S_c$, total savings, $\Delta s = s_c - s_w > 0$ and where, to simplify, we have assumed that the function f_1 in (12) is additive.

It is then not too difficult to show (see Sordi, 2001, p. 101) that the MVM1 reduces to the following *3D*-dynamical system in *E*, *U*, and *X*

$$\dot{E} = \left[\frac{s_w}{\overline{v}} - g_n + \frac{\Delta s}{\overline{v}}X\left(1-U\right)\right]E \tag{13}$$

$$\dot{U} = \left[h\left(E\right) + \delta\hat{E} - \alpha\right]U \tag{14}$$

$$\dot{X} = \left[\frac{\Delta s}{\overline{v}} - \frac{s_c}{\overline{v}}U - \frac{\Delta s}{\overline{v}}X\left(1-U\right)\right]X \tag{15}$$

with positive singular point $(E^*, U^*, X^*) = (E^*, 1 - \overline{v}\, g_n/s_c, s_c(g_n - s_w/\overline{v})/\Delta s g_n)$ where E^* is the value of the employment rate for which $h(E^*) = \alpha$.

It is worth noting that the positive equilibrium guarantees *steady-state results* that have a Pasinettian–Kaldorian 'flavour':

- it guarantees a steady-state growth of the system at a warranted rate equal to the natural rate and is such that the *Cambridge equation* $r^* = (1 - U^*)/\overline{v} = g_n/s_c$ is satisfied;
- in order to be economically meaningful, it requires that the Pasinettian case holds

$$0 \le s_w < \overline{v}g_n < s_c \le 1 \tag{16}$$

- it is such that the steady-state growth path is characterized by a positive (constant) rate of unemployment equal to $(1 - E^*)$ rather than by full employment.

However, and more importantly given our purposes, when condition (16) is satisfied, the system may not converge to (E^*, U^*, X^*), but *rather persistently fluctuate around it* (see Figure 15.7).[13]

15.4. A HYBRID VERSION OF THE GROWTH CYCLE MODEL

The MVM1, apart from the generalized Phillips curve and the assumption concerning differential savings, maintains all the other simplifying assumptions of the OVM. In particular, it assumes a permanent product market equilibrium and does not have an independent investment function. In the attempt to integrate the two different types of dynamic model considered in this chapter ('Keynesian' and 'Marxian'),[14] we next propose a second modified version of the model (MVM2), in which (*i*) investment in fixed capital is explained by an 'accelerator-type' mechanism;[15] and (*ii*) the product market does not clear at all times; rather it is governed by a simplified version of the error-adjustment mechanism (3) of the 'multiplier–accelerator' model studied in Section 15.2.

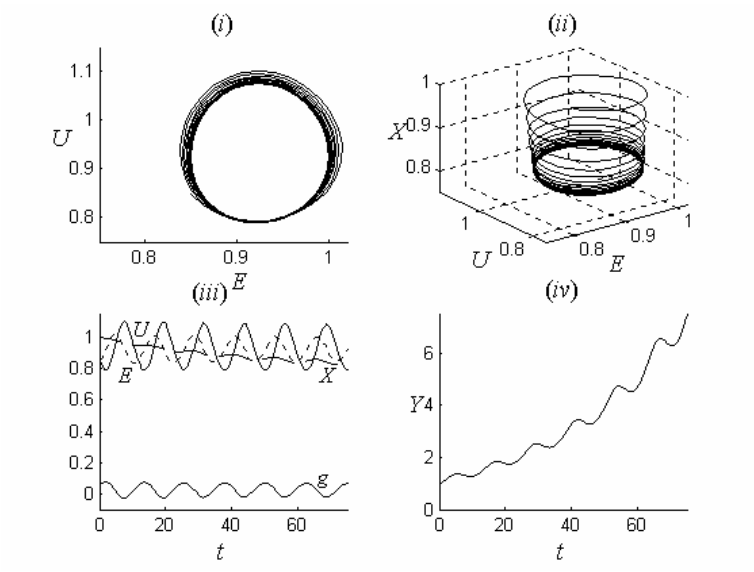


Figure 15.7: Growth cycles of the MVM1 (for $s_c = 0.9$, $s_w = 0.0134$, $\bar{v} = 2.57$, $\alpha = 0.0221$, $\beta = 0.0037$, $\gamma = 0.9$, $\rho = 1$, $\delta = 0.02$ and $X(0) = 1$)

First of all, with regard to induced investment,[16] we slightly modify (4) by assuming that the desired level of the capital stock K^d is determined by the flexible accelerator (see Goodwin, 1948), expressed in terms of *expected* output Y^e:[17]

$$I = \dot{K} = K^d - K = \bar{v}Y^e - K \tag{17}$$

Assuming that Y^e is related to realized output according to an *extrapolative* mechanism (see Gandolfo, 1997, p. 210), we can write

$$Y^e = Y + \tau\dot{Y}, \ \tau > 0$$

from which, inserting into (17), we obtain

$$I = \left(\bar{v} + \bar{v}\tau g_Y - v\right)Y \tag{18}$$

where $g_Y = \hat{Y}$ and $v = K/Y$.

Second, we consider the following error-adjustment mechanism for disequilibrium in the product market[18]

$$\dot{Y} = g_n Y + \eta\left(I - S\right), \ \eta > 0 \tag{19}$$

where, as in the OVM, we assume that $s_w = 1$, which implies:

$$K_w = 0 \leftrightarrow X = 1$$

$$S = s_c\left(1 - U\right) Y$$

Thus, expressing (19) in terms of rates of growth

$$g_Y = \frac{g_n}{1 - \eta\bar{v}\tau} + \frac{\eta}{1 - \eta\bar{v}\tau}\left[\bar{v} - v - s_c\left(1 - U\right)\right] \tag{20}$$

Taking account of (18) and (20), it is then not too difficult to show that the MVM2 reduces to the following complete 3D-dynamical system in the three endogenous variables *E*, *U* and *v*[19]

$$\dot{E} = \left\{\frac{\eta\bar{v}\tau g_n}{1 - \eta\bar{v}\tau} + \frac{\eta}{1 - \eta\bar{v}\tau}\left[\bar{v} - v - s_c\left(1 - U\right)\right]\right\}E \tag{21}$$

$$\dot{U} = \left[h\left(E\right) + \delta\hat{E} - \alpha\right] U = \left[\ F\left(E, U, v\right) - \alpha\ \right] U \tag{22}$$

$$\dot{v} = \bar{v} - v + \left(\tau\bar{v} - v\right)\left\{\frac{g_n}{1 - \eta\bar{v}\tau} + \frac{\eta}{1 - \eta\bar{v}\tau}\left[\bar{v} - v - s_c\left(1 - U\right)\right]\right\} \tag{23}$$

where the function $F(E, U, v)$ in equation (22) is such that $F_E(E, U, v) = h'(E) > 0$, $F_U(E, U, v) = \delta\eta s_c/(1 - \eta\tau\bar{v})$, $F_v(E, U, v) = -\delta\eta/(1 - \eta\tau\bar{v})$.

The positive equilibrium point for this version of the model proves to be (E^{**}, U^{**}, v^{**}), where, as in the MVM1, $E^{**} = E^*$, whereas we now have

$$U^{**} = 1 - \frac{g_n v^{**}}{s_c}, \ v^{**} = \frac{\left(1 + \tau g_n\right)\bar{v}}{\left(1 + g_n\right)}$$

It is worth noting that, as in the MVM1, the equilibrium values just obtained guarantee interesting *steady-state results*. In particular

- they guarantee a steady-state growth of output and capital stock at a rate equal to the natural rate

$$g_Y^{**} = \frac{g_n}{1 - \eta\bar{v}\tau} + \frac{\eta}{1 - \eta\bar{v}\tau}\left[\bar{v} - v^{**} - s_c\left(1 - U^{**}\right)\right] = g_n$$

$$g_K^{**} = \frac{\bar{v} + \bar{v}\tau g_Y^{**} - v^{**}}{v^{**}} = g_n$$

- in order to be economically meaningful, they require that the Pasinettian case $0 < g_n v^{**} < s_c \le 1$ holds;

- they imply the Cambridge equation:

$$r^{**} = \frac{1-U^{**}}{v^{**}} = \frac{g_n v^{**}}{s_c}\frac{1}{v^{**}} = \frac{g_n}{s_c}$$

However, as in the MVM1, the system may not converge to the positive equilibrium point, but rather persistently fluctuate around it. The next two sections are devoted to the analysis of the dynamics of the MVM2, from both a qualitative and numerical point of view.

15.4.1. Qualitative Analysis of the Dynamics

Linearising the dynamical system (21)–(23) at (E^*, U^{**}, v^{**}), we obtain

$$\begin{bmatrix} \dot{E} \\ \dot{U} \\ \dot{v} \end{bmatrix} = \mathbf{J}^{**} \begin{bmatrix} E-E^* \\ U-U^{**} \\ v-v^{**} \end{bmatrix} = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} E-E^* \\ U-U^{**} \\ v-v^{**} \end{bmatrix}$$

where

$$a_{12} = \frac{\partial \dot{E}}{\partial U} = \frac{\eta s_c E^*}{1-\eta\bar{v}\tau},\ a_{13} = \frac{\partial \dot{E}}{\partial v} = -\frac{\eta E^*}{1-\eta\bar{v}\tau}$$

$$a_{21} = \frac{\partial \dot{U}}{\partial E} = h'\left(E^*\right)\ U^{**} > 0,\ a_{22} = \frac{\partial \dot{U}}{\partial U} = \frac{\delta\eta s_c U^{**}}{1-\eta\bar{v}\tau},\ a_{23} = \frac{\partial \dot{U}}{\partial v} = -\frac{\delta\eta U^{**}}{1-\eta\bar{v}\tau}$$

$$a_{32} = \frac{\partial \dot{v}}{\partial U} = \frac{\left(\tau\bar{v}-v^{**}\right)\ \eta s_c}{1-\eta\bar{v}\tau},\ a_{33} = \frac{\partial \dot{v}}{\partial v} = -\left(1+g_n\right) - \frac{\left(\tau\bar{v}-v^{**}\right)\ \eta}{1-\eta\bar{v}\tau}$$

Thus, the characteristic equation of the linearised system is

$$\lambda^3 + A\lambda^2 + B\lambda + C = 0 \tag{24}$$

where

$$A = \frac{-\eta v^{**} - \delta\eta s_c U^{**} + 1 + g_n - \eta\tau\bar{v}g_n}{1-\eta\bar{v}\tau} \tag{25}$$

$$B = -\frac{\eta s_c U^{**}}{1-\eta\bar{v}\tau}\left[\delta\left(1+g_n\right) + E^* h'\left(E^*\right)\right] \tag{26}$$

$$C = -\frac{h'\left(E^*\right)U^{**}\eta s_c E^*\left(1+g_n\right)}{1-\eta\bar{v}\tau} \tag{27}$$

Under the following two assumptions:

Assumption 1: $1-\eta\tau\bar{v} < 0 \leftrightarrow \tau > \dfrac{1}{\eta\bar{v}} > 0$

Assumption 2: $1-\eta v^{**} - \delta\eta s_c U^{**} + \left(1-\eta\tau\bar{v}\right)g_n < 0$

from (25), (26) and (27) it follows that all coefficients of (24) are positive so that the sign of the expression $AB - C$ is undetermined. In particular, in terms of the parameters of the model, it is easy to check that we have $AB - C \gtreqless 0$ according as to whether

$$F\left(\tau\right) = -\left[\left(\tau\bar{v}-v^{**}\right) - \delta s_c U^{**}\right]\left[\delta\eta\left(1+g_n\right) + \eta E^* h'(E^*)\right]$$

$$-\delta\left(1-\eta\bar{v}\tau\right)\left(1+g_n\right)^2 \gtreqless 0 \tag{28}$$

We are then in a position to prove the following proposition:

Proposition 1: *Under Assumptions 1 and 2, if expectations are extrapolative and such that*

$$\frac{1}{\eta\bar{v}} < \tau \leq \frac{v^{**}}{\bar{v}} = \frac{(1+\tau g_n)}{(1+g_n)} \leftrightarrow \frac{1}{\eta\bar{v}} < \tau \leq 1 \tag{29}$$

the positive equilibrium (E^*, U^{**}, v^{**}) of the dynamical system of the MVM2 is locally stable.

Proof: We already noted that, under Assumptions 1 and 2, all coefficients of the characteristic equation (24) are positive. Moreover, when (29) holds, from (28) it follows that we also have

$$F(\tau) > 0 \leftrightarrow AB - C > 0$$

All Routh–Hurwitz conditions for (local) stability of the positive equilibrium point are therefore satisfied. ■

However, when condition (29) does not hold and the only restrictions on the value of τ are those implied by Assumptions 1 and 2, *we can establish the possibility of persistent cyclical paths* of the variables of the model. More precisely, we can establish the following result:

Proposition 2: *Under Assumptions 1 and 2, there exists a value of the parameter $\tau = \tau_H > 1$ at which the dynamical system (21)–(23) of the MVM2 undergoes a Hopf bifurcation.*

Proof: It is easy to adjust to the present case the method of proof of the Hopf bifurcation adopted in previous contributions, for example in Asada (1995). To this end, let us note that, under Assumptions 1 and 2, all coefficients of the characteristic equation (24) are positive. Moreover, from (28) we know that $AB - C \gtreqless 0$ according to whether $F(\tau) \gtreqless 0$. Proposition 1 implies that $AB - C > 0$ for all values of $\tau \leq 1$. Let us instead indicate with $\bar{\tau} > 1$ a value of the parameter for which $F(\bar{\tau})<0$ so that $A(\bar{\tau})B(\bar{\tau})-C(\bar{\tau})<0$ and the equilibrium is locally unstable. By continuity, this means that there exists at least one value of the parameter $\tau=\tau_H \in (1,\bar{\tau})$ at which

$$A(\tau_H)B(\tau_H)-C(\tau_H)=0 \tag{30}$$

$$\left.\frac{\partial\left[A(\tau)B(\tau)-C(\tau)\right]}{\partial\tau}\right|_{\tau=\tau_H} \neq 0 \tag{31}$$

Now, it is possible to prove (see Asada, 1995, p. 248) that, under Assumption 1, (30) is a *necessary and sufficient* condition for (24) to have a pair of purely imaginary roots $\lambda_{1,2} = \pm i\omega$ $(\omega \neq 0)$. Then it is also straightforward to prove (see Asada, 1995, pp. 267–8) that condition (31) implies that the real parts of the complex roots for $\tau = \tau_H$ cross the real axis at non-zero speed. We can therefore conclude that the dynamical system (21)–(23) undergoes a Hopf bifurcation at $\tau = \tau_H$. ∎

15.4.2. Numerical Simulations

We finally turn to the illustration of the analytical findings of Section 15.4.1 with the help of numerical simulations.

Cyclical convergence to the positive equilibrium. As a starting point of our numerical analysis, we choose $\eta = 1$, $s_c = 0.9$, $\bar{v} = 2.57$, $\tau = 0.75$, $\alpha = 0.0221$, $\beta = 0.0037$, $\gamma = 0.9$, $\rho = 1$ and $\delta = 0.02$ which imply $(E^*, U^{**}, v^{**}) = (0.9221,$ $0.9268, 2.5538)$ and are such that Assumptions 1 and 2 and condition (29) of Proposition 1 are satisfied.

As was to be expected, the trajectory (starting, for example, from initial conditions equal to $E(0) = 0.81$, $U(0) = 0.81$ and $v(0) = 2.4$) converges to the positive equilibrium (see Figures 15.8(*i*) and 15.8(*ii*)). The convergence appears to be cyclical, so that we can conclude that the equilibrium (E^*, U^{**}, v^{**}) is a *stable focus*. This implies that, after a certain period of time, when the growth cycle has completely dampened out, output starts to grow exponentially at a rate equal to the natural rate $g_n = 0.0258$ (see Figures 15.8(*iii*) and 15.8(*iv*)).

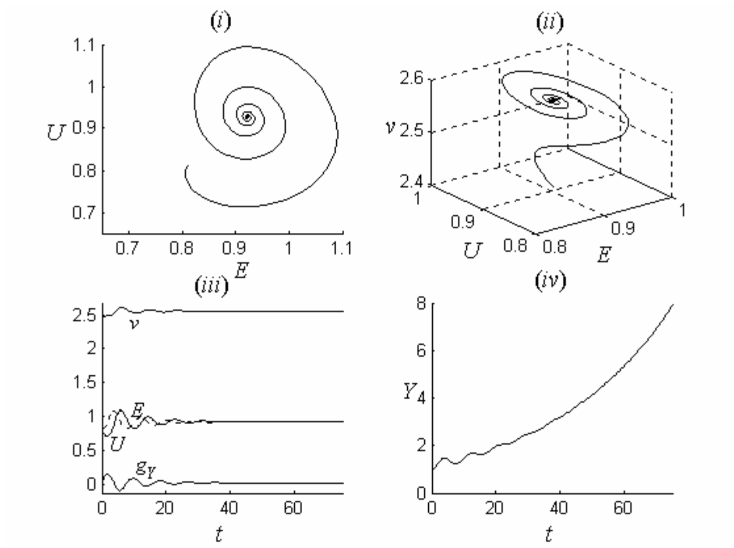


Figure 15.8: Locally stable positive equilibrium of the MVM2

Convergence to a stable limit cycle. Proposition 2 proves only the *existence part* of the Hopf bifurcation theorem and says nothing about the uniqueness and stability of the closed orbits. The numerical simulations we have performed for a large number of initial conditions, however, seem to suggest that the emerging closed orbit is stable (i.e., the Hopf bifurcation *super-critical*) and unique. For example, with initial conditions $E(0) = 0.91$, $U(0) = 0.91$ and $v(0) = 2.4$ and the same set of parameter values as before (except that $\tau = 1.0211$, such that Assumptions 1 and 2 but not condition (29) are satisfied), the trajectory converges to a limit cycle (see Figures 15.9(*i*) and 15.9(*ii*)). Thus, in this case, output is subject to *persistent cyclical growth* (see Figures 15.9(*iii*) and 15.9(*iv*)).

15.5. CONCLUSIONS

In the wake of the recent spurt of interest in the study of the interaction between economic growth and economic fluctuations, in this chapter we

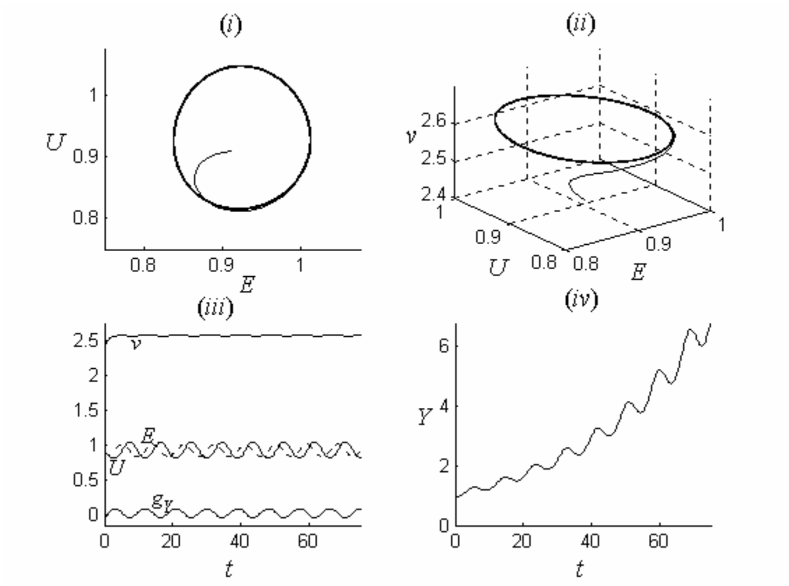


Figure 15.9: Stable limit cycle of the MVM2

presented two prototype models belonging to the non-market clearing approach to macrodynamics, and analysed them with regard to their capacity to generate cyclical growth paths of the relevant variables.

First of all, we considered the simple Keynesian ‘multiplier–accelerator’ model which, while providing basic insights into the role played in the dynamics of the economy by effective demand problems, can generate cyclical growth paths only when an exogenous trend is superimposed on it. We showed that the version of the model with a non-linear formulation of the accelerator can generate cyclical growth paths also in the case in which we assume that the autonomous (‘innovational’) component of investment is a periodic function of time. This, as a first approximation, can be interpreted as a way of introducing into the model Schumpeter’s idea of clustering of innovations. As a second prototype model, we then presented Goodwin’s model which, by contrast, provides a neat explanation of growth cycles based on the Marxian reserve army mechanism. The OVM is known to be structurally unstable and it focuses only on the labour market, in sharp contrast with the other prototype model. In the attempt to overcome these two drawbacks of the OVM, we proposed two extensions of it and showed that both modified versions of the model (MVM1 and MVM2) entail a three-dimensional dynamical system. In both cases, then, by application of the Hopf bifurcation theorem, we were able to prove that the system can admit

limit cycle solutions. Thus, both overcome the first of the two drawbacks of the OVM mentioned above.

It is worth recalling that the increase in the dimensionality of the dynamical system in the MVM1 is simply due to the introduction into the OVM of the hypothesis of differential savings along Kaldorian–Pasinettian lines, together with a more general specification of the Phillips curve: the rest of the OVM is left unchanged. In the MVM2, on the other hand, such an increase is due to the introduction into the OVM of an independent investment function (which corresponds to the flexible accelerator), together with the assumption of disequilibrium also in the product market. In this sense, the MVM2 represents an attempt at integrating the two prototype models we considered in the paper in such a way that also the second of the two shortcomings of the OVM is overcome. In performing this integration, however, we considered a simplified version of the Keynesian prototype model. It remains to be ascertained whether the integrated model can generate cyclical growth paths in the case in which a more general specification of investment behaviour is used. By this we mean a specification which, besides an error-adjustment mechanism for induced investment, takes account of the interpretation of the autonomous component of investment outlined in Section 15.2.2, and tries to develop it. This issue, given the recent revival of Schumpeterian ideas within the endogenous growth approach, appears to be worth investigating further.

NOTES

1. This was stressed, in many occasions, by Goodwin. See, for example, Goodwin (1986). See also Kaldor (1954, pp. 53–6).
2. One should be aware that, more recently, in the last two decades or so, there has been a renewed interest in the problem of the interrelation of economic growth and business cycles, within both the literature on real business cycles (e.g., Kydland and Prescott, 1982) and the literature on endogenous growth (see Aghion and Howitt, 1998, ch. 8). This more recent (*market clearing*) literature is critically surveyed by Fiaschi and Sordi in their contribution to this volume.
3. This feature of linear models is one of the most critically debated by the ‘pioneers’ of economic dynamic modelling such as Goodwin (e.g., 1953, 1955), Harrod (e.g., 1939, 1951), Kaldor (1954) and Kalecki (e.g., 1968). See also Pasinetti (1960) and Allen (1967, ch. 18 and 19).
4. In order to guarantee continuity with what follows, we consider a continuous-time formulation of the model. That is, rather than Samuelson’s (1939) original formulation or its extension by Hicks (1950), we consider the version of the model studied by Phillips (1954). As usual, in what follows a dot over a variable (e.g., $\dot{z}$) indicates its derivative with respect to time (dz/dt), whereas a hat (e.g., $\hat{z}$) indicates its rate of growth ($\dot{z}/z$).
5. In presenting these results, we are simply adjusting to our case and notation the presentation of Phillips’s model made by Allen in his celebrated book on Macroeconomic Theory (1967,

pp. 328–33). In short, given the characteristic equation of (5), $\lambda^2 + (1/\varepsilon)(\varepsilon + s - \overline{v})\lambda + (s/\varepsilon) = 0$, and the expression for its roots, $\lambda_{1,2} = (1/2\varepsilon)[(\overline{v} - \varepsilon - s) \pm \Delta^{1/2}]$, where $\Delta = (\overline{v} - \varepsilon - s)^2 - 4s\varepsilon$, the intervals of parameter values listed in Table 15.1 are obtained by jointly considering the conditions under which these roots are real or complex ($\Delta \geq 0$ or $\Delta < 0$) and the stability condition ($\overline{v} < \varepsilon + s$). On the application of this kind of qualitative analysis to Phillips's model, see also Flaschel (1993, pp. 100–01).

6. In passing, it should be noted that this (with $\overline{v}$ *exactly* equal to $\varepsilon + s$) is the only case in which the simple linear multiplier–accelerator model we are considering is able to represent oscillations which last in time, neither exploding, nor dying away.
7. Mathematically, the cyclical component $Y^c(t)$ is given by the general solution of the homogenous equation, whereas the growth or trend component $Y^p(t)$ is given by the particular solution of the non-homogenous equation.
8. On this type of criticism, see Kaldor (1954). See also Aghion and Howitt (1998, p. 234).
9. See Goodwin (1951, p. 9). In the numerical simulations which follow, we will use the following functional form for ϕ:

$$\phi(\dot{Y}) = \left\{ \frac{I^{\max} + I^{\min}}{I^{\max} \exp[-(I^{\max} + I^{\min})\overline{v}\dot{Y} / I^{\max} I^{\min}] + I^{\min}} - 1 \right\} I^{\min}$$

which, as required by the nonlinear accelerator, is such that $\phi(0) = 0$, $\phi'(0) = \overline{v}$, $\phi(-\infty) = -I^{\min} < 0$ and $\phi(+\infty) = I^{\max} > 0$. For a similar specification, see Allen (1967, pp. 378–80).
10. As stressed by Gandolfo (1997, p. 522), forced oscillators of this kind have recently attracted the attention of students of economic dynamics for their capacity to generate transient chaotic dynamics. Of course, given an appropriate choice of parameter values, this is also the case for equation (7) with (9).
11. In Aghion's and Howitt's opinion (1998, p. 234), this is perhaps the first model in which the occurrence of economic fluctuations was modelled as a deterministic consequence of the accumulation (i.e., growth) process; more specifically, of the variations in income distribution this process induces over time.
12. To the best of my knowledge, this more general version of the Phillips curve was first introduced in the OVM by Cugno and Montrucchio (1982, p. 97).
13. See Sordi (2001), where this is proved by applying to the dynamical system (13)–(15) the Hopf bifurcation theorem. For the numerical simulation, we have used a linear approximation of the generalized Phillips Curve (12), of the type $f_1(E, \hat{E}) \approx -\gamma + \rho E + \delta \hat{E}$.
14. In doing this, we are following a suggestion made by Flaschel, Franke and Semmler (1997, p. 101), who maintain that 'a proper integration of a Goodwin growth cycle with the multiplier–accelerator and the study of the dynamics arising from it is still an open problem and should be addressed by future research'.
15. For other contributions in which an independent investment function has been introduced into the 'growth–cycle' framework see, for example, Glombowski and Krüger (1988), Rampa and Rampa (1988) and Wolfstetter (1982). See also Flaschel (1988, 1993).
16. In order to keep things simple, we assume that $I_a(t) = 0$ for all t.
17. As is easy to verify, (17) and (4) are equivalent when $Y^e = Y$ and $\theta = 1$.
18. For a similar assumption made in a different extension of Goodwin's model, see Glombowski and Krüger (1988, p. 427).
19. The derivation of this and other calculations in the paper are available from the author upon request.

16. Real business cycle models, endogenous growth models and cyclical growth: a critical survey

Davide Fiaschi and Serena Sordi

16.1. INTRODUCTION

Since the early contributions to the topic, business cycles have been considered as essentially connected with the development of capitalist economies. In the early 1980s the issue of the relationship between growth and cycles was addressed within a market clearing environment by the real business cycle (RBC) literature, and, more recently, in endogenous growth (EG) models.[1]

In the articles belonging to the RBC literature, which are based on the neoclassical model of (optimal) capital accumulation augmented by technology shocks, it is common to find assertions like 'our approach integrates growth and business cycle theory' (Kydland and Prescott, 1982, p. 1345) or 'real business cycles theory (...) holds considerable promise for enhancing our understanding of economic fluctuations and growth as well as their interaction' (King, Plosser and Rebelo, 1988, p. 196). Thus, it appears that such an integration of growth and business cycle theory is understood as one of the most, if not *the* most, important achievement of the analysis.

The relationship between growth and cycles has also been tackled from a different point of view; some contributions in the EG literature focus on the possibility of generating non-linear (periodic) dynamics as the effect of introducing endogenous growth into an otherwise neoclassical growth model.

In the whole of this variegated literature on RBC and EG, one can distinguish at least *three main different approaches* to the study of the interaction between growth and cycles:

1. Starting with the contributions by Kydland and Prescott (1982) and Long and Plosser (1983), RBC theorists have studied the interaction between growth and cycles within a *stochastic* business cycles framework, where cycles are generated by continuous exogenous shocks to technology;

2. Some EG theorists analyse the implications of a sounder microfoundation of technological progress on the relationship between growth and cycles. Examples are Aghion and Saint-Paul (1998a), where cycles are generated in a model with a Schumpeterian flavour, and Stadler (1990), where growth is generated by a learning-by-doing process;
3. Finally, the possibility of multiple steady states in EG models has given importance to the analysis of out-of-steady-state dynamics (cycles) in a deterministic framework (see Greiner and Semmler, 1996a, 1996b; and Benhabib and Perli, 1994).

The aim of this chapter is to survey and compare these contributions, starting from the stochastic approaches 1 and 2 (Sections 16.2) and continuing with the deterministic approach 3 (Section 16.3). Section 16.4 concludes and gives some suggestions for further research.

16.2. GROWTH MODELS WITH STOCHASTIC BUSINESS CYCLES

The first attempts to explain business cycles on the basis of stochastic shocks are due to Frisch (1933) and Slutsky (1937). While the latter showed how the sum of random components generates cycles similar to empirical fluctuations, the former presented technical innovations as exogenous perturbations to the available level of technological progress. In the same period Schumpeter (1939) identified in the continuous introduction of new innovations, and the resulting shocks to productivity, the source of growth of a country. In this view growth and business cycle are generated by the same source and therefore they must be jointly analysed. Unlike Frisch, however, Schumpeter considered innovations as driven by economic factors, i.e., *as endogenous*.[2] These contributions are the main inspiration of the modern theory of RBC, where the business cycle is seen as a phenomenon essentially due to shocks to the real part of the economy and long-run growth as the cumulative sum of such shocks.

The seminal contribution by Kydland and Prescott (1982) starts from the idea of analysing – within the neoclassical framework of optimising agents – the behaviour of an economy which is converging toward its long-run equilibrium but which is continually shocked by random disturbances. The standard RBC model is essentially a neoclassical growth model in which exogenous technological progress is modelled as a stochastic process (see Cooley and Prescott, 1995). The aim is to simulate series whose properties are similar to those of observed series. Their conclusion is that the business cycle is essentially a real phenomenon and that the neoclassical framework

can account for most of the cyclical component. This conclusion, however, can be criticised from many points of view. We focus on the assumption of exogenous technological change. Stadler (1990) and Aghion and Saint-Paul (1998a) provide two interesting contributions, whose findings substantially differ from RBC results. The next section is devoted to the exposition of a standard RBC model, which then will be used as a term of comparison for the two models of endogenous technological progress presented in Section 16.2.2.

16.2.1. A Basic Model of RBC

In this section we present a basic RBC model, referring, in particular, to the classical contribution by Christiano and Eichenbaum (1992), in which the authors analyse an RBC model with government expenditure and endogenous labour supply. For the sake of simplicity, in our presentation, we ignore the government expenditure.

The model economy considered is composed by an infinitely living representative agent, which maximises the discount sum of instantaneous utility. To keep things simple we consider a log instantaneous utility function, $U = \ln c_t + \phi V(l_t)$, where c_t is the per capita consumption and l_t is leisure at period *t*.

The production side of the economy is characterised by a competitive market where each firm produces homogeneous output according to a Cobb–Douglas constant returns to scale technology:

$$y_t = \left(A_t z_t h_t\right)^{1-\theta} k_t^{\theta} \tag{1}$$

where y_t is per capita output, A_t, an index of long-run deterministic technological progress, z_t, an index of short-run cyclical productivity, $h_t = L - l_t$ are the worked hours, and k_t is per capita capital.

Then it is assumed that productivity evolves according to the following stochastic process[3]

$$z_t = z_{t-1} \exp\left(\gamma + \varepsilon_t\right) \tag{2}$$

where $\gamma > 0$ is a constant drift and $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$.

From the assumption of competitive markets it follows that factors are paid according to their marginal productivity, so that

$$r_t = \theta\left[\left(A_t z_t h_t\right)^{1-\theta} k_t^{\theta-1}\right] - \delta$$
$$w_t = (1-\theta)\left[A_t^{1-\theta} z_t^{1-\theta} h_t^{-\theta} k_t^{\theta}\right]$$

where δ is the depreciation rate of capital.

In this economy the first welfare theorem holds, so that it is convenient to solve the competitive allocation as a social planner problem, that is[4]

$$\max_{\{c_t, l_t\}_{t=0}^{\infty}} W = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \ln c_t + \phi V(l_t) \right\} \tag{3}$$

subject to

$$k_{t+1} = (Az_t h_t)^{1-\theta} k_t^{\theta} - c_t + (1-\delta) k_t$$
$$k_0 = k(0)$$

Since z_t follows a stochastic path, a closed form solution to problem (3) does not exist in general, but only for a particular configuration of the parameters ($\theta = \delta = 1$ and $V(l_t) = \ln l_t$). The procedure generally used consists in calculating the steady-state equilibrium (which corresponds to the locus which the economy converges to if it were not subject to productivity shocks) and in approximating problem (3) around this steady state.[5] In order to calculate the steady-state equilibrium, it is useful to normalise each growing variable with respect to its long-run growth rate. By so doing, we obtain the following normalised variables:

$$\hat{k}_t = \frac{k_t}{A_{t-1} z_{t-1}}, \ \hat{c}_t = \frac{c_t}{A_t z_t}, \ \hat{y}_t = \frac{y_t}{A_t z_t}, \ \hat{z}_t = \frac{z_t}{z_{t-1}}$$

The solution to problem (3) is given by the following difference equations

$$h_{t+1} = q_h \hat{k}_t^{n_h} \hat{z}_t^{m_h}; \ \hat{k}_{t+1} = q_k \hat{k}_t^{n_k} \hat{z}_t^{m_k} \tag{4}$$

where the coefficients q_i, n_i and m_i for $i = h, m$ are non-linear functions of the original parameters of the model.

The system of difference equations (3) – together with the stochastic process (2) – fully describes the dynamics of model. As already stressed, a closed form solution does not exist and q_i, n_i and m_i must be determined by numerical simulations. Christiano and Eichenbaum (1992, p. 441) provides a numerical simulation for which per capita capital shows a stochastic trend (the average growth rate of which equals to the value of the drift of stochastic process). On the contrary, hours worked do not show any trend. Further the system proves to be stable.

The properties of the artificial model economy are analysed by means of numerical simulations. In so doing, the purpose is to match some empirical regularities or *stylised facts* (see Canova, 1998). These regularities generally refer to differences in the variance of some relevant variables and/or cross-correlations among the latter.

Let us turn our attention to the problem of the choice of proper indicators for business cycles. We start from the standard deviations of the variables. In

Table 16.1 we report the standard deviations of our simulated series after they are logged and detrended by the Hodridk–Prescott filter to extract the transitory component at frequency 4–6 years and, for comparison, we also report the corresponding statistics relative to US data (see Canova, 1998).[6] As we see, the simulated series show standard deviations normalised with respect to the standard deviation of output that in some cases overestimate the real values (consumption and productivity) while in others underestimate them (investment, hours worked and real wages).

Table 16.1 – Standard deviations, sources US estimates: Canova (1998)

Statistic of σ	*Simulated data*	*US data 1955:3–1986:3*
σ_c/σ_y	0.6174	0.49
σ_i/σ_y	2.2055	2.82
σ_h/σ_y	0.3710	1.06
σ_w/σ_y	0.6785	0.70
$\sigma_{y/n}/\sigma_y$	0.6949	0.49

Another point is the correct sign and magnitude in the cross-correlation among variables. In Table 16.2 we report such cross-correlation for simulated series (firstly detrended and then filtered by the Hodrick–Prescott procedure) and for US data. Table 16.2 shows the main drawback of the RBC model, namely, the discordance between simulated and real series with respect to the cross-correlation of hours worked and productivity with output. The negative empirical correlation between hours worked and productivity suggests that the labour market does not work as a perfectly competitive market. Moreover, simulations predict a strong correlation between productivity and output whereas the empirical evidence shows only a slightly positive correlation. According to Canova (1998, p. 503) this correlation changes over time so that there is 'the need for theoretical work to provide reasons for why this phenomenon occurs'.

Table 16.2 – Correlations, sources US estimates: Canova (1998)

Statistic of σ	Simulated data	US data 1955:3–1986:3
corr(c, y)	0.9481	0.75
corr(I, y)	0.9677	0.91
corr(h, y)	0.8890	0.88
corr(w, y)	0.9705	0.81
corr(y/h, y)	0.9684	0.10
corr(h, w)	0.7691	0.67
corr(h, y/h)	0.7887	–0.24

The latter point suggests that the current RBC theory needs a sounder microfoundation of technological progress. This will be the argument of the next section.

16.2.2. Microfoundation of Technological Progress

The exogeneity of technological progress was the most serious criticism faced by exogenous growth models *à la* Solow. In the same way some authors argue that a business cycle theory needs a microfoundation of innovation. Moreover, as we showed in the previous section, the RBC model predicts a strong positive correlation between productivity and output, which contrasts with the empirical evidence.[7] Thus, endogenising technological progress may help to understand the causes of this phenomenon.

Firstly, we analyse the model proposed by Stadler (1990), in which technological progress is viewed as a *learning-by-doing* process, i.e., as a by-product of production. In so doing, our purpose is to stress that, even if productivity is always strongly procyclical, this has crucial implications for business cycles analysis, in particular for the effect of monetary policy on long-run productivity.

Secondly, we consider the contribution by Aghion and Saint-Paul (1998b), in which technological progress is endogenised within a Schumpeterian framework. The focus in this case is on the relationship between productivity and the growth rate of output; it will be shown that productivity can be both procyclical and countercyclical.

Stadler (1990) proposes a model in which growth is generated only through *learning by doing*. Business cycles are generated by shocks to productivity similar to those in the model presented above but without a positive drift. In the model economy there is no fixed capital and the only accumulated factor is knowledge, which is accumulated through learning by doing. The learning-by-doing process is external to the firm and is a by-product of production. In such a framework, the consumer plays no role in the allocation of resources. As a consequence, the condition for maximisation of profits of the representative firm determines the competitive allocation.

The representative firm solves the following problem:

$$V = \max_{\{h_t\}_{t=0}^{\infty}} E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[y_t - \left(\frac{w_t}{p_t} \right) h_t \right] \right\} \tag{5}$$

where $0 < \beta < 1$ is a discount factor,[8] w_t, the monetary wage, p_t, the price level and h_t, the labour employed by the firm.

It is assumed that the production function is:

$$y_t = z^{\chi} h_t^{1-\theta} a_t^{\eta}, \text{ with } 0 < \theta < 1, \chi > 0, \eta > 0 \tag{6}$$

where z_t is the productivity shock and a_t the per capita accumulated knowledge available to the firm. The latter is accumulated both by an exogenous component and as a by-product of production:[9]

$$a_t = a_{t-1}^{1-\delta_a} y_{t-1}^{\lambda} \tag{7}$$

where λ, $\delta_a > 0$.

The productivity shock follows a stochastic process similar to (2) but without drift:

$$z_t = z_{t-1} \exp(\varepsilon_t) \quad \varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right) \tag{8}$$

Given the assumption of exogenous knowledge accumulation, the intertemporal maximization problem (5) of the representative firm becomes a static optimisation problem. Therefore the demand curve of the firm is given by the first order condition (FOC) to the problem of the maximisation of profits at time *t*, that is:

$$h_t^d = \left[\frac{(1-\theta) z_t^{\chi} a_t^{\eta}}{w_t / p_t} \right]^{1/\theta} \tag{9}$$

Labour supply is assumed to be[10]

$$h_t^s = \exp(\phi_1) \left(\frac{w_t}{p_t} \right)^{\phi_2} \quad \text{with } \phi_1, \phi_2 > 0 \tag{10}$$

Stadler (1990) then assumes that neither consumer nor firm knows the level of price at period *t* when they bargain in the labour market to set the wage at the beginning of the period. Thus, they formulate their choice according to their expectations of the level of prices at time *t*, given their information set (which included the level of productivity z_{t-1}).

In our case, we obtain the following expression for the log of aggregate supply $\tilde{y}_t^s$

$$\tilde{y}_t^s = \bar{f}_0 + \bar{f}_1 \tilde{z}_t + \bar{f}_2 \tilde{z}_{t-1} + \bar{f}_3 \tilde{a}_t + \bar{f}_4 \left(\tilde{p}_t - \tilde{p}_t^e \right) \tag{11}$$

where '~' over a variable denotes the log of the variable, $\tilde{f}_0 = (1-\theta)[\phi_2 \log(1-\theta) + \phi_1]/(1+\phi_2\theta)$, $\tilde{f}_1 = \chi/\theta$, $\tilde{f}_2 = -(1-\theta)\chi/\theta(1+\phi_2\theta)$, $\tilde{f}_3 = (1-\theta)\phi_2\eta/(1+\phi_2\theta)$ and $\tilde{f}_4 = (1-\theta)/\theta$. Hence aggregate supply depends on productivity shocks, $\tilde{z}_t$, accumulated knowledge, $\tilde{a}_t$, and on the difference between actual price and expected price $\tilde{p}_t - \tilde{p}_t^e$; it is to be noted that an expected increase in demand can lead to an increase in supply of output. We next deal with this point.

On the basis of the quantitative theory, following Stadler (1990) we assume that aggregate demand is given by

$$y_t^d = \frac{m_t}{p_t} \tag{12}$$

where m_t is per capita money stock. m_t is assumed not to be observable at the beginning of period t but follows the stochastic process

$$m_t = m_{t-1}\exp\left(\gamma_m + \zeta_t\right) \quad \zeta_t \sim N\left(0, \sigma_\zeta^2\right) \tag{13}$$

where $\gamma_m > 0$ is the drift in money supply.

In equilibrium we have:

$$\tilde{p}_t - \tilde{p}_t^e = \frac{\zeta_t - \tilde{f}_1\varepsilon_t}{1+\tilde{f}_4} \tag{14}$$

Substituting (14) in (11) yields equilibrium (log) output:

$$\tilde{y}_t = q_0 + q_1\tilde{z}_{t-1} + q_2\tilde{a}_t + q_3\zeta_t + q_4\varepsilon_t \tag{15}$$

where $q_0 = [\phi_2\log(1-\theta) + \phi_1]/(1+\theta\phi_2\})$, $q_1=\chi(1+\phi_2)/(1+\phi_2\theta)$, $q_2=\eta(1+\phi_2)/(1+\phi_2\theta)$, $q_3 = 1-\theta$ and $q_4=\chi$.

The dynamics of output is driven by monetary and real shocks (ζ and ε, respectively), by productivity shocks, $\tilde{z}$, and by the stock of accumulated knowledge, $\tilde{a}$. Equation (15), provided that $\tilde{a}_t$ is time-constant, makes it clear why monetary shocks cannot have a long-run effect in an RBC model: in fact, only the level of productivity $\tilde{z}$ is relevant in determining the long-run behaviour of $\tilde{y}$, while the monetary shock ζ has only a short-run effect. However, if $\tilde{a}$ incorporates also the short-run dynamics of $\tilde{y}$, as in our model, then monetary shocks can have a long-run effect.

Substituting (7) and (8) in (15) leads to:

$$\begin{aligned}\tilde{y}_t =& q_0 + q_1\sum_{i=1}^{t-1}\varepsilon_i + q_2\left\{\left(1-\delta_a+\lambda q_2\right)^t\tilde{a}_0 + \lambda\sum_{j=0}^{t}\left(1-\delta_a+\lambda q_2\right)^{t-j}\cdot\right.\\ &\left.\left[q_0 + q_1\sum_{i=1}^{j-1}\varepsilon_i + q_3\zeta_j + q_4\varepsilon_j\right]\right\} + q_3\zeta_t + q_4\varepsilon_t\end{aligned} \tag{16}$$

From equation (16) we can calculate the expected level of long-run (log) output:

$$E_0\left[\tilde{y}_t\right] = q_0\left(\frac{\delta_a}{\delta_a-\lambda q_2}\right) + q_2\left(\tilde{a}_0 + \frac{\lambda q_0}{\lambda q_2-\delta_a}\right)\left(1-\delta_a+\lambda q_2\right)^t$$

The existence of a positive long-run growth of output depends on the value of $1 - \delta_a + \lambda q_2$. If $1 - \delta_a + \lambda q_2 > 1$, then the growth rate of output is ever increasing; otherwise, the long-run growth rate is zero and the level of output (in log) converges to $q_0\delta_a/(\delta_a - \lambda q_2)$.

From equation (17) it follows that if the economy achieves positive long-run growth, i.e. $1 - \delta_a + \lambda q_2 > 1$, then both monetary and real shocks have long-run effects. However they work through two different channels. Monetary shocks affect the long-run output through the stock of accumulated knowledge, while real shocks work both directly on output and indirectly on the stock of accumulated knowledge. In this regard assume $\eta = 0$, i.e. the stock of accumulated knowledge is not relevant in production; this implies that $q_2 = 0$ (the long-run growth rate is equal to zero). From (16) we have:

$$\tilde{y}_t = q_0 + q_1\sum_{i=1}^{t-1}\varepsilon_i + q_3\zeta_t + q_4\varepsilon_t$$

Therefore real shocks ε have long-run effects, while monetary shocks ζ have only one-period effects. This case corresponds to the standard RBC model of Section 16.2.1, provided that the stochastic process of productivity has a positive drift.

Finally, it is to be noted that monetary shocks alone can account for the unit root detected in empirical data. Consider the case where real shocks have no effect on output, that is $\chi = 0$, from which $q_1 = q_4 = 0$. From (16) we have

$$\begin{aligned}\tilde{y}_t =& q_0\left[1+\lambda\sum_{j=0}^{t}\left(1-\delta_a+\lambda q_2\right)^{t-j}\right] + q_2\left(1-\delta_a+\lambda q_2\right)^t\tilde{a}_0\\ &+\lambda q_3\sum_{j=0}^{t}\left(1-\delta_a+\lambda q_2\right)^{t-j}\zeta_j + q_3\zeta_t\end{aligned}$$

and, since $1 - \delta_a + \lambda q_2 > 1$, monetary shocks ζ have long-run effects and $\tilde{y}_t$ presents a unit root (or more than a unit root in our case where growth rate is ever increasing).

16.2.2.1. Endogenous growth theory and the Schumpeterian approach to innovation

The relationship between productivity and growth is hotly debated in the literature. According to Schumpeter, recessions are required to eliminate inefficient firms from the market, so that the final effect of a decrease in economic activity is an increase in overall productivity (see Caballero and Hammour, 1991). There are however factors suggesting that productivity is procyclical, like learning by doing, demands spillovers and capital market imperfections that constrain investment in the R&D sector (see Stiglitz, 1993). Moreover, Bean (1990) stresses that, if the reorganization of the firm is costly in terms of output, during a recession the opportunity cost of such an activity is lower and therefore the procyclical pattern of productivity is magnified. The empirical evidence for the relationship between productivity and output is mixed. Canova (1998), for example, shows that productivity was countercyclical up to the mid 1960s and procyclical afterwards.

In what follows we discuss a simplified version of the Aghion and Saint-Paul (1998b) model, which provides an example of how productivity can be countercyclical if innovations are endogenised in a Schumpeterian fashion. The basic idea is that the cost of innovation is lower in recession than in a boom if the act of innovating drains resources from production. This framework corresponds to the case analysed by Hall (1999), who assumes that the increases in productivity are the result of internal reorganization of a firm, which negatively affects the current level of output.[11]

In the economy at period t there are N_t firms and N_t different goods; each firm is a monopolist in its own market. Each is characterised by a level of productivity $z_{i,t}$ and it is assumed that firms cannot change their employed workforce and production capacity. Thus, the gross product of each firm is determined only by its productivity, which is given by $\exp(z_{i,t})$.

Let $v_{i,t} = dz_{i,t}/dt$ be the change in productivity. Each firm can modify its productivity by sacrificing part of its production; a change in productivity equal to $v_{i,t}$ involves a proportional drop of output equal to $\mu = \mu(v_{i,t})$, where the authors assume that $\mu(0) = 0$, $\mu' > 0$, $\mu'' > 0$.

For firm i, demand at period t is given by:

$$d_{i,t} = \left(\frac{D_t}{P_t}\right)\left(\frac{p_{i,t}}{P_{\ t}}\right)^{-\psi} \tag{17}$$

where $\psi > 1$, D_t is an aggregate demand index and P_t, an aggregate price index, defined by

$$P_t = \left(\int_0^{N_t} p_{i,t}^{1-\psi}\, di\right)^{1/(1-\psi)} \tag{18}$$

The net output of firm i, $y_{i,t}$, is given by:

$$y_{i,t} = \left[1-\mu\left(v_{i,t}\right)\right]\exp\left(z_{i,t}\right) \tag{19}$$

Given that the equilibrium in the monopolist market i is given by $y_{i,t} = d_{i,t}$, we then have:

$$p_{i,t} = \left[1-\mu\left(v_{i,t}\right)\right]^{-1/\psi}\exp\left(-\frac{z_{i,t}}{\psi}\right)Y_t^{1/\psi}P_t^{(\psi-1)/\psi} \tag{20}$$

where $p_{i,t}$ is the price of good i produced by firm i.

Given the actual level of productivity, firm i must choose the optimal increase in productivity; the latter is the result of the following problem:[12]

$$V\left(z_{i,t}\right) = \max_{v_i}\ \left\{\pi_{i,t}dt + \left(1-rdt\right)E_t\left[V\left(z_{i,t}+v_i dt\right)\right]\right\} \tag{21}$$

where V is the value of the firm, r is the constant interest rate at which agents can lend or rent their resources. The future value of the firm depends on the level of prices, which, in turn, depend on the aggregate demand index; the expectation operator reflects the possible uncertainty of the latter variable.

To determine the equilibrium, market entry and exit conditions must be specified. Following Aghion and Saint-Paul (1998b) we assume that the firm must bear a fixed cost equal to C to enter the market, while the firm has a liquidation value equal to τC, where $\tau < 1$.[13] Finally, it is assumed that new firms show the same level of productivity. In the equilibrium the number of firms will be constant if the expected value of a firm V is greater than τC, but lower than C, that is

$$\dot{N} = 0 \leftrightarrow V \in \left[\tau C, C\right] \tag{22}$$

In equilibrium we expect that $V \in [\tau C, C]$ since V greater than C drives new firms to enter the market, causing V to decrease up to C, while a firm’s value lower than τC leads firms to exit, causing V to increase up to τC.

Consider the symmetric equilibrium where $p_{i,t} = p_t$ and $z_{i,t} = z_t$ and therefore $v_{i,t} = v_t$ for all i.[14] From (22) we obtain the level of profits for each firm:

$$\pi_t = \frac{D_t}{N_t} = d_t \tag{23}$$

where $d_t = D_t/N_t$ is the demand for the single firm.

Aghion and Saint-Paul (1998b) assume that cycles are generated by fluctuations in the level of aggregate demand D_t. In particular, it is assumed that there exist only two states, E, expansion and R, recession, and that the probability of jumping from state R to E follows a Poisson process and is given by ε, while from state E to R it is given by ζ.

In the stochastic steady state for each possible state E and R, all variables, except for z and p, are constant and the economy stays for a fraction of time equal to $\varepsilon/(\zeta+\varepsilon)$ in expansion and for a fraction equal to $\zeta/(\zeta+\varepsilon)$ in recession. In such a framework the problem of the firm (21) becomes:[15]

$$\begin{cases} V_t^R = \max_v \left\{\pi_t^R dt + (1-rdt)\left[(1-\varepsilon dt)V_{t+dt}^R + \varepsilon dt V_{t+dt}^E\right]\right\} \\ V_t^E = \max_v \left\{\pi_t^E dt + (1-rdt)\left[(1-\zeta dt)V_{t+dt}^E + \zeta dt V_{t+dt}^R\right]\right\} \end{cases} \tag{24}$$

where $\pi_t^j = D_t^j / N_t = d_t^j$ for $j = E, R$.

To close the model we must calculate the level of *per capita* demand in the two states; in fact, since the number of firms is endogenous, a higher level of aggregate demand may not be matched by a higher level of per capita demand.

We assume that recession is deep enough to lead some firms to exit; in this case the free entry condition implies that

$$V_t^R = \tau C;\ V^E = C \tag{25}$$

By inserting (25) into problem (24), we obtain the levels of profits for each of the two states:

$$\pi^R = C\left[r\tau - \varepsilon(1-\tau)\right];\ \pi^E = C\left[r + \zeta(1-\tau)\right]$$

from which the levels of per capita demand also easily follow (see equation (23)):

$$d^R = C\left[r\tau - \varepsilon(1-\tau)\right];\ d^E = C\left[r + \zeta(1-\tau)\right] \tag{26}$$

From these expressions, it follows that $d^E > d^R$, i.e. per capita demand is higher in expansion than in recession.

From equations (26) and the FOCs of problem (24), we obtain:

$$\begin{aligned} \frac{\mu'(v^R)}{1-\mu(v^R)} &= \frac{\tau}{r\tau - \varepsilon(1-\tau)} \\ \frac{\mu'(v^E)}{1-\mu(v^E)} &= \frac{1}{r + \zeta(1-\tau)} \end{aligned} \tag{27}$$

which provide implicit solutions for v^R and v^E. From (27), given the assumptions on μ, it follows that $v^R > v^E$: this means that productivity in this model is countercyclical. The intuition is straightforward: the opportunity cost of innovating, measured in terms of lower output, is higher in expansion because the level of demand (and therefore of profits) is greater than in recession. This is the main finding of the model: recession can have a positive impact on the growth rate of productivity. Finally, from (27) it follows that $\partial v^j/\partial r < 0$, $j = E, R$, $\partial v^R/\partial \varepsilon > 0$, $\partial v^R/\partial \zeta < 0$, $\partial v^R/\partial \tau < 0$ and $\partial v^E/\partial \tau > 0$.

The business cycle affects the growth rate of the economy because of the different increases in productivity which characterise the two states. The average growth rate of the economy is given by:

$$g = \left(\frac{\zeta}{\zeta+\varepsilon}\right)v^R + \left(\frac{\varepsilon}{\zeta+\varepsilon}\right)v^E \tag{28}$$

With regard to the average growth rate of the economy, Aghion and Saint-Paul (1998b, p. 333) underline the importance of the following three effects:

1. *Composition effect*: the time that the economy spends in expansion, $\varepsilon/(\zeta + \varepsilon)$, and in recession, $\zeta/(\zeta+\varepsilon)$, given $v^R > v^E$, crucially affects the average growth rate;
2. *Return effect*: in expansion per capita demand is higher than in recession and therefore the longer the economy spends in expansion, the higher are the increases in productivity;
3. *Cost of capital effect*: since the firm does not recoup all costs of entry in the case of exit, both the time it spends in expansion and in recession and the liquidation value affect the incentive to increase productivity.

Given all this, we can conclude that the three main parameters (ε, ζ, τ) of the model affect the average growth rate in the following way:

- $\partial g/\partial \varepsilon > 0$ if $dv^R/d\varepsilon > (v^R - v^E)/(\zeta + \varepsilon)$; in the case of an increase in ε, the composition effect is negative $(\zeta(v^E - v^R)/(\zeta+\varepsilon)^2 < 0)$, whereas the sum of the return and of the cost of capital effects is positive $([\zeta/(\zeta + \varepsilon)]/(dv^R/d\varepsilon) > 0)$, so that $\partial g/\partial \varepsilon > 0$ if the latter is greater than the former.
- $\partial g/\partial \zeta < 0$ if $-dv^E/d\zeta > (v^R - v^E)/(\zeta+\varepsilon)$; in the case of an increase in ζ the composition effect is positive $\varepsilon(v^E - v^R)/(\zeta+\varepsilon)^2 > 0)$, whereas the sum of the return and of the cost of capital effects are negative $([\varepsilon/(\zeta + \varepsilon)](dv^E/d\zeta) < 0)$, so that $\partial g/\partial \zeta < 0$ if the former is greater than the latter.
- $\partial g/\partial \tau > 0$ if $dv^E/d\tau > -(\zeta/\varepsilon)(dv^R/d\tau)$; the cost of capital effect is positive for v^E $([\varepsilon/(\zeta + \varepsilon)](dv^E/d\tau) > 0)$ and negative for $v^R([\zeta/(\zeta\ \varepsilon)]dv^R/d\tau) < 0)$, so that $\partial g/\partial \tau > 0$ if the former is greater than the latter.

To sum up, the idea that expansion is better that recession for growth is challenged in this model; our intuition is based on the fact that during a recession the reorganization of a firm in order to increase its productivity is less costly; this has implications for both the cross-correlation between output and productivity (which is negative rather than positive) and the effect of recession on average growth rate (which is positive rather than negative).

16.3. ENDOGENOUS GROWTH MODELS WITH DETERMINISTIC CYCLES

Since the late 1980s, following seminal contributions by Lucas (1988) and Romer (1986a, 1990), a large number of articles have focused on the dynamics of EG models, in a *deterministic* context. This line of research was strongly motivated by the fact that the original contributions by Romer and Lucas focused on steady state only and, in addition, neglected the stability properties of the steady state.

We now turn to the analysis of this literature. Our main purpose is to check whether there exist results concerning the emergence of persistent cycles similar to those obtained by Benhabib and Nishimura (1979) for the conventional exogenous growth models. A positive answer to this question would imply that EG models offer an additional (deterministic) approach to the study of cyclical growth in a market-clearing context, alternative to the stochastic approach presented in Section 16.2.

In reviewing this literature on EG, it is useful to distinguish between one-sector models and two-sector models. In models belonging to the first class (e.g., Romer, 1986a), the accumulation of knowledge is only a by-product of production activities and EG is generated by mechanisms of learning by doing, by externalities or by increasing returns. The models belonging to the second class, on the other hand, starting either from Lucas (1988) or from Romer (1990), generate EG by assuming an intentional allocation of resources for the accumulation of human capital or an intentional R&D effort for increasing the level of technological progress.

In what follows, we first review the existing literature on this topic with regard to one-sector EG models (Section 16.3.1). We will focus in particular on recent contributions by Greiner and Semmler (e.g., 1996a, 1996b), which aim to show that a basic model of EG with learning by doing (which is a modified version of the Romer 1986a model) may produce a rich array of outcomes, such as multiple steady states, indeterminacy of equilibria or even persistent cycles of the state variables. Then, in Section 16.3.2 we tackle the same problem with regard to two-sector EG models. In this case, we discuss the classical contribution by Benhabib and Perli (1994) and we briefly

present the results of a recent contribution by Mattana and Venturi (1999), in which they show that periodic solutions may emerge in the Lucas model.

16.3.1. Persistent Cycles in One-Sector EG Models

The model considered by Greiner and Semmler (1996a, 1996b) is a one-sector EG model of the Romer type (with learning by doing) in which, however, it is assumed that one unit of investment has different effects concerning the building up of physical capital and knowledge. This implies that the two variables cannot be merged into a single variable.

The production possibilities of the model economy (in per capita terms) are given by

$$\frac{Y}{L} = bA^{\alpha}k^{1-\alpha},\ b > 0,\ \alpha \in (0,1) \tag{29}$$

where A stands for the stock of knowledge, K, the stock of physical capital, L, labour force, $k = K/L$. In what follows, to simplify, we choose $b = 1$.

Assuming that L grows exponentially in time at a constant rate equal to $n > 0$, the equation for the evolution of k is the following:

$$\dot{k} = i - (\delta + n)k$$

where $\delta > 0$ is the rate of depreciation, $i = I/L$ and I, gross investment.

With regard to the stock of knowledge, it is assumed that it accumulates according to a *learning-by-doing* process *à la* Arrow, in the formulation given by Levhari (1966). In addition, it is assumed that the contribution – to the formation of knowledge – of gross investment further back in time is smaller than that of recent gross investment. Hence (see Greiner and Semmler, 1996a, p. 82):

$$A(t) = \rho \int_{-\infty}^{t} \exp\left[\rho(s-t)\right] i(s)\,ds,\ \rho > 0 \tag{30}$$

or

$$\dot{A} = \rho(i - A) = \rho\left(A^{\alpha}k^{1-\alpha} - c - A\right) \tag{31}$$

where $\rho > 0$ represents the weight given to more recent levels of gross investment and c stands for per capita consumption.[16]

Following Greiner and Semmler (1996a, p.82), we limit ourselves to analysing the competitive situation, in which the evolution of knowledge is not explicitly taken into account by the representative agent when solving the optimisation problem. Normalising so as to have $L(0) = 1$, the latter is the following:[17]

$$\underset{c}{\text{Max}} \int_0^\infty \exp\left[-(\beta-n)t\right] u(c(t))\,dt,\ \beta > 0 \tag{32}$$

subject to: $\dot{k} = A^{\alpha} k^{1-\alpha} - c - (\delta+n)k$.

From the current-value Hamiltonian for problem (32), we then obtain the following set of necessary FOCs for an optimum:[18]

$$u'(c) - \lambda = 0 \leftrightarrow u'(c) = \lambda \tag{33}$$

$$\dot{\lambda} = \lambda(\delta+\beta) - \lambda(1-\alpha)\left(\frac{A}{k}\right)^{\alpha} \tag{34}$$

We are now in a position to derive the differential equations system which describes the dynamics of the model economy. From (33)–(34), one obtains

$$\left[\frac{u''(c)c}{u'(c)}\right]\frac{\dot{c}}{c} = -\sigma\frac{\dot{c}}{c} = (\delta+\beta) - (1-\alpha)\left(\frac{A}{k}\right)^{\alpha}$$

which, together with (31) and (33), gives

$$\frac{\dot{c}}{c} = \frac{(1-\alpha)}{\sigma}\left(\frac{A}{k}\right)^{\alpha} - \frac{\delta+\beta}{\sigma} \tag{35}$$

$$\frac{\dot{k}}{k} = \left(\frac{A}{k}\right)^{\alpha} - \left(\frac{c}{k}\right) - (\delta+n) \tag{36}$$

$$\frac{\dot{A}}{A} = \rho\left(\frac{A}{k}\right)^{\alpha-1} - \rho\left(\frac{c}{A}\right) - \rho \tag{37}$$

Following a standard practice, the order of system (35)–(37) can be reduced by performing a change of variables with $k_A = k/A$ and $c_A = c/A$. Hence:

$$\frac{\dot{k}_A}{k_A} = k_A^{-\alpha} - \frac{c_A}{k_A} - (\delta+n) - \rho k_A^{1-\alpha} + \rho(1+c_A) \tag{38}$$

$$\frac{\dot{c}_A}{c_A} = \frac{(1-\alpha)}{\sigma} k_A^{-\alpha} - \frac{(\delta+\beta)}{\sigma} - \rho k_A^{1-\alpha} + \rho(1+c_A) \tag{39}$$

A balanced growth path for the original system is obtained as a rest point of the reduced system at which $\dot{k}_A/k_A = \dot{c}_A/c_A = 0$ so that $\dot{A}/A = \dot{k}/k = \dot{c}/c$.

As shown by the authors (see *Proposition 2* in Greiner and Semmler 1996a, p. 85 and *Theorem 1* and *Theorem 2* in Greiner and Semmler 1996b, p. 110), in the case of positive per capita growth,

(*i*) if $\delta+n \geq (\delta+\beta)/\sigma$, there exists a unique steady state (c_A^*, k_A^*) which proves to be saddle stable and such that:

$$c_A^* = \frac{k_A^{*1-\alpha} - \rho k_A^{*2-\alpha} - (\delta+n)k_A^* + \rho k_A^*}{1-\rho k_A^*}$$

$$-\frac{(1-\alpha)}{\sigma}(\rho k_A^* - 1) + k_A^*\left[\rho - \frac{(\delta+\beta)}{\sigma}\right] + \rho k_A^{*1+\alpha}\left[-(\delta+n) + \frac{(\delta+\beta)}{\sigma}\right] = 0$$

(*ii*) if $\delta+n < (\delta+\beta)/\sigma$, there exist two steady states, (c_{A1}^*, k_{A1}^*) and (c_{A2}^*, k_{A2}^*) with $k_{A1}^* < k_{A2}^*$, such that the path associated with the second can be anything except a saddle path.

Given our aim, we are mainly interested in case (*ii*) in which the second of the two steady states is either completely stable or unstable. To understand why it is so, let us note that complete stability requires that the Jacobian of the linearised system at (c_{A2}^*, k_{A2}^*)

$$\mathbf{J}_2^* = \begin{bmatrix} -\alpha k_{A2}^{*-\alpha} - (1-\alpha)\rho k_{A2}^{*1-\alpha} + c_{A2}^*/k_{A2}^* & \rho k_{A2}^* - 1 \\ -\left[\alpha(1-\alpha)/\sigma\right] c_{A2}^* k_{A2}^{*-\alpha-1} - (1-\alpha)\rho c_{A2}^* k_{A2}^{*-\alpha} & \rho c_{A2}^* \end{bmatrix}$$

is such that

$$\text{tr}\,\mathbf{J}_2^* = -\alpha k_{A2}^{*-\alpha} - (1-\alpha)\rho k_{A2}^{*1-\alpha} + \frac{c_{A2}^*}{k_{A2}^*} + \rho c_{A2}^* < 0$$

$$\det \mathbf{J}_2^* = \left[-\alpha k_{A2}^{*-\alpha} - (1-\alpha)\rho k_{A2}^{*1-\alpha} + \frac{c_{A2}^*}{k_{A2}^*}\right]\rho c_{A2}^* + (\rho k_{A2}^* - 1)\left[\frac{\alpha(1-\alpha)}{\sigma} c_{A2}^* k_{A2}^{*-\alpha-1} + (1-\alpha)\rho c_{A2}^* k_{A2}^{*-\alpha}\right] > 0$$

The basic fact is that, as one of the parameters, e.g. ρ, varies, there may exist a value ρ_H for which $\text{tr}\,\mathbf{J}_2^*(\rho_H) = 0$ and $\det \mathbf{J}_2^*(\rho_H) > 0$. When this happens, the dynamics of the system may undergo a qualitative change, known in the literature as *Hopf bifurcation*:[19] *the model economy does not reach the steady-state growth rate, but rather persistently fluctuates around it.*

In terms of the parameters of the model, we have:[20]

$$\mathrm{tr}\mathbf{J}_2^* = 0 \leftrightarrow c_{A2}^* = \frac{\alpha k_{A2}^{*-\alpha} + (1-\alpha)\rho_H k_{A2}^{*1-\alpha}}{\left(\rho_H + k_{A2}^{*-1}\right)} \tag{40}$$

$$\begin{aligned} \det \mathbf{J}_2^* > 0 \leftrightarrow c_{A2}^* &= \frac{\alpha k_{A2}^{*-\alpha} + (1-\alpha)\rho_H k_{A2}^{*1-\alpha}}{\left(\rho_H + k_{A2}^{*-1}\right)} \\ &> \frac{\alpha(1-\alpha)(1-\rho_H k_{A2}^*) + \rho_H \sigma k_{A2}^*}{\rho_H \sigma k_{A2}^*} \\ \leftrightarrow -(1-\alpha-\sigma)\alpha\rho_H^2 k_{A2}^{*2} &+ \rho_H \sigma k_{A2}^* (1-\alpha) + \alpha(1-\alpha) < 0 \end{aligned} \tag{41}$$

The existence of a Hopf bifurcation also requires that

$$\left.\frac{d}{d\rho}\mathrm{tr}\mathbf{J}_2^*(\rho)\right|_{\rho=\rho_H} \neq 0 \tag{42}$$

Using parameter values that satisfy conditions (41) and (42), numerical simulations show that indeed the model may generate limit cycles which, in addition, prove to be stable (see Greiner and Semmler, 1996a, pp. 91–96 and 1996, pp. 111–16). The conclusion is that the model may generate persistent cycles in the growth rate of per capita variables, e.g. per capita output, namely, persistent *growth cycles*.

16.3.2. Persistent Cycles in Two-Sector EG Models

It is a fact that, with few exceptions of the kind considered in the previous subsection, the contributions on out-of-steady-state dynamics in EG models have focussed on two-sector models, for both the case with human capital (see, for example, Caballé and Santos, 1993, Mulligan and Sala-i-Martin, 1993, Benhabib and Perli, 1994, Boldrin and Rustichini, 1994, Xie, 1994, Barro and Sala-i-Martin, 1995, ch. 5, and Arnold, 1997) and with R&D (see, for example, Benhabib, Perli and Xie, 1994, Asada, Semmler and Novak, 1998, and Arnold, 2000a, 2000b). The result is a large body of literature in which it is possible to find some clear-cut results. First, no interesting out-of-steady-state dynamics is usually found for the *social planner solution* of the models. This is shown, for example, by Caballé and Santos (1993) who find saddle-path stability for the Lucas model in the absence of externalities (see also Barro and Sala-i-Martin, 1995 and Arnold, 1997), whereas Asada, Semmler and Novak (1998) obtain the same result for the social planner problem of the Romer model. Second, starting with the important contributions by Benhabib and Perli (1994) and Xie (1994), interesting dynamics (included multiple equilibria, indeterminacy of equilibria and even the emergence of *periodic solutions*) has been shown for the market solution of the Lucas model when externalities from human capital are considered. Given the aim of the present chapter, we now turn to a brief description and analysis of this second set of results.

In the model by Lucas (1988) the optimisation problem that has to be solved by the representative agent is the following

$$\max_{c(t),u(t)} \int_0^\infty \left(\frac{c^{1-\sigma}-1}{1-\sigma}\right)\exp(-\beta t)\,dt \tag{43}$$

subject to: $\dot{k} = k^{1-\alpha}(uh)^\alpha h_a^\gamma - c \quad \dot{h} = \delta h(1-u)$

$$k(0) = k_0 > 0,\ h(0) = h_0 > 0$$

where c is per capita consumption, β, a positive discount factor, σ, the inverse of the intertemporal elasticity of substitution, k, physical capital, h, human capital, u, the fraction of labour allocated to the production of physical capital (so that uh is the fraction of effective labour), δ, a positive technology parameter, $1-\alpha$, the share of capital, and γ, a positive externality parameter in the production of human capital.

For problem (43) we obtain the following set of necessary FOCs for an interior solution:[21]

$$c^{-\sigma} = \lambda_1 \tag{44}$$

$$\frac{\alpha c^{-\sigma} k^{1-\alpha} h^{\alpha+\gamma-1} u^{-\alpha}}{\delta} = \lambda_2 \tag{45}$$

$$\dot{\lambda}_1 = \beta\lambda_1 - \lambda_1(1-\alpha)k^{-\alpha}h^{\alpha+\gamma}u^\alpha \tag{46}$$

$$\dot{\lambda}_2 = \beta\lambda_2 - \lambda_1 \alpha k^{1-\alpha}h^{\alpha+\gamma-1}u^\alpha - \lambda_2\delta(1-u) \tag{47}$$

whereas the transversality condition is[22]

$$\lim_{t\to\infty}\exp(-\beta t)[\lambda_1 k + \lambda_2 h] = 0$$

From (44)–(47), with simple algebraic manipulation, we derive the dynamical system of the model. In order to do so, note that from (44) and (46) we obtain:

$$\dot{c} = \left(\frac{1-\alpha}{\sigma}\right) c k^{-\alpha} h^{\alpha+\gamma} u^{\alpha} - \left(\frac{\beta}{\sigma}\right) c \tag{48}$$

whereas from (45) and (47)

$$-\sigma c^{-1}\dot{c} + (1-\alpha)k^{-1}\dot{k} + (\alpha+\gamma-1)h^{-1}\dot{h} - (1-\alpha)u^{-1}\dot{u} = \beta - \delta$$

Inserting in the latter expression (48) and the constraints of (43), we finally obtain

$$\dot{x}_2 = \eta x_2^2 + \frac{\alpha\psi}{1-\alpha} x_2 - x_2 x_3$$

where $x_2 = u$, $x_3 = c/k$, $\eta = (1 - \alpha - \gamma)\delta/(1 - \alpha)$, $\psi = -(\alpha + \gamma)\delta/\alpha < 0$.

We then define

$$x_1 = h^{-(\alpha+\gamma)/\alpha} k$$

from which, taking the derivative with respect to time of both sides and inserting the two constraints of (43)

$$\dot{x}_1 = \psi(1-x_2)x_1 - x_1 x_3 + x_1 x_2^{\alpha}$$

Finally

$$\dot{x}_3 = x_3^2 + \phi x_3 x_1^{-\alpha} x_2^{\alpha} - \xi x_3$$

where $\phi = \left[(1-\alpha)/\sigma\right] - 1$ and $\xi = \beta/\sigma$.

Thus the results of Benhabib and Perli (1994) can be discussed by analysing the following *reduced* dynamical system

$$\dot{x}_1 = x_1^{1-\alpha} x_2^{\alpha} + \psi(1-x_2)x_1 - x_1 x_3 \tag{49}$$

$$\dot{x}_2 = \eta x_2^2 - \left(\frac{\psi\alpha}{\phi}\right) x_2 - x_2 x_3 \tag{50}$$

$$\dot{x}_3 = x_3^2 + \phi x_1^{-\alpha} x_2^{\alpha} x_3 - \xi x_3 \tag{51}$$

the steady-state values of which are[23]

$$x_1^* = \left[\frac{(1-\alpha)\xi - \delta(\alpha+\gamma) - \delta(1-\alpha-\gamma)x_2^*}{(1-\alpha)\phi}\right]^{-1/\alpha} x_2^*$$

$$x_2^* = 1 - \frac{\alpha(\beta-\delta)}{\delta\left[\gamma - \sigma(\alpha+\gamma)\right]}, \quad x_3^* = \eta x_2^* + \frac{\delta(\alpha+\gamma)}{1-\alpha}$$

The study of the local stability of the linearised system leads to some major results that can be summarised as follows (see *Propositions 1* and *2* in Benhabib and Perli, 1994, p. 123–24)

1. For values of the parameters such that $0 < \beta < \delta$ and $\sigma > (1 + \beta/\psi)$, the Jacobian of the system has one eigenvalue with a negative real part and two eigenvalues with a positive real part. In this case the competitive equilibrium path is locally unique: given the initial conditions of *k* and *h*, there exists only one value of *u* and *c* that drives the economy towards its (determinate) steady-state path;
2. For some other values, namely for $\delta < \beta < -\psi$, $\sigma < (1 + \beta/\psi)$ and $\gamma > 1 - \alpha$, one eigenvalue has a positive real part and two eigenvalues have a negative real part. In this case there is always a continuum of equilibria in the neighbourhood of the indeterminate steady-state path: given the initial conditions of *k* and *h*, there exist infinitely many values of *u* and *c* that drive the economy towards the BGP;
3. When β and σ are as in the previous case, but $0 < \gamma \leq 1 - \alpha$, there is either (*i*) one positive eigenvalue and two eigenvalues with negative real parts or (*ii*) three eigenvalues with positive real parts. In the latter case there is *complete* instability and no equilibrium path leads to the BGP.

As stressed by Benhabib and Perli (1994, p. 124), of particularly interest is Case 3, in which there is a basic change in the roots structure of the Jacobian. This happens when, as one of the parameters (for example, the externality parameter γ) changes, the real parts of the two complex eigenvalues change sign. At the value $\gamma = \gamma_H$ for which the real parts of the two complex eigenvalues are zero the system may undergo a Hopf bifurcation such that it persistently fluctuates around the steady-state path. However, although this possibility is mentioned by Benhabib and Perli (1994, p. 124), they do not investigate the matter any further but rather concentrate on the indeterminacy results.

An interesting analysis of the case is instead contained in a recent contribution by Mattana and Venturi (1999), where the authors, after a general presentation of Benhabib and Perli’s results, concentrate on the possibility of emergence of periodic solutions and achieve results worth mentioning. They establish analytically the existence of closed orbits (see *Theorem 1* in Mattana and Venturi, 1999, p. 270) for the reduced system (46)–(48). Moreover (see *Theorem 2*), they establish, with the help of numerical simulation, that the closed orbits emerging from the steady state can be either sub-critical (i.e., repelling) or super-critical (i.e., attracting). In particular, when σ is 'small', the numerical simulations show that the supercritical seems to prevail.

16.4. CONCLUSIONS

In this paper we explored the possibility of generating persistent growth cycles in a market-clearing framework, distinguishing between stochastic and deterministic models. As regards the first class of models, we focussed on the novelty originating from the new growth theory. In particular, we stressed both the role played by monetary factors and the consequences of modelling R&D activity. With regard to the second class, we analysed the conditions for the emergence of periodic orbits in EG models which, in such a framework, represents growth cycles.

The inverse empirical relationship between productivity and output is a major challenge for business-cycle researchers. The representative agent framework, which is common to all contributions considered in this chapter, does not appear well-suited to this goal. As suggested by Lippi (1993), this may require the consideration of the dynamics of the productivity of firms and of their interaction.

Finally, analysis of out-of-steady-state dynamics in EG models has shown interesting results, but the latter appear to be applied to a very narrow set of EG models. A direction for future research is to explore the possibility of periodic orbits in other types of deterministic EG models, such as R&D Schumpeterian models, in an attempt to bridge the gap between stochastic and deterministic approaches.

NOTES

1. In another contribution to this volume, Sordi analyses the problem of the interrelation between growth and cycle in a non-market clearing framework.
2. Frisch (1933) argued that his theory of the business cycle was supported by the Schumpeterian idea of innovations as a cause of economic fluctuations. However, he thought that Schumpeter was considering as exogenous the process of innovation, while Schumpeter was actually distinguishing between *scientific discovery*, not driven by economic forces, and *innovation*, i.e. the economic implementation of a scientific discovery, which depends on economic factors.
3. The form of this stochastic process is suggested by the empirical evidence about the presence of a unit root in time series.
4. We assume that population is constant. In the case in which the population increases at a constant rate equal to n, the representative agent's problem becomes:

$$\max_{\{c_t, l_t\}_{t=0}^{\infty}} W = \sum_{t=0}^{\infty} \beta^t (1+n)^t \{\ln c_t + \phi V(l_t)\} dt$$

subject to the intertemporal budget constraint

$$k_{t+1} = (A_t z_t h_t)^{1-\theta} k_t^{\theta} - c_t + (1-\delta-n) k_t$$

In this case, writing $\hat{\beta} = \beta(1+n)$ and $\hat{\delta} = \delta + n$, the problem becomes equivalent to (3).
5. The details of such a procedure are in an appendix available upon request.
6. The purpose in filtering simulated series is to make them comparable with the actual time series, that is to consider cycles of the same frequencies for both.
7. As shown by Christiano and Eichenbaum (1992), this result holds for a wide range of parameters.
8. β depends on the interest rate and, in a model with capital accumulation, would be an endogenous variable; however, in the framework considered by Stadler, without fixed capital, the constancy of β does not appear a restrictive assumption.
9. To avoid scale effects, which are typical in endogenous growth models, we suppose that every firm has a positive externality proportional to per capita capital knowledge stock available in the economy.
10. A possible microfoundation of such labour supply can be performed by considering a representative agent that maximises an instantaneous utility function whose arguments are consumption and leisure; moreover labour is assumed to be paid to the real wage rate *w*/*p*, which represents the only source of income for the agent. The choice of the labour supply (10) is due to its analytical tractability. We also notice that the total amount of labour which is supplied has no upper bound; a more plausible formulation would entail total available labour being fixed and finite. For example a possible formulation for the labour supply would be $h_t^s = [1+(w_t/p_t)^{-v}]^{-1} L$, which is the solution to the following consumer problem: $U = c_t^v (L - h_t)^v$, s. to $c_t = (w_t/p_t) h_t$.
11. The increase in productivity, however, could also be the result of employing in the production process new resources brought onto the market (see Stiglitz, 1993). We analyse this more general case in an Appendix available on the authors' websites, in which the firm can choose between these two methods of increasing its productivity.
12. Problem (21) is derived from the standard Bellman equation for a firm whose goal is to maximizes its value, given by the discounted sum of future expected profits at rate *r*.
13. Aghion and Saint-Paul (1998b) highlight how the presence of a liquidation value independent of productivity can generate an exit effect, which leads firms to underinvesting; for the sake of simplicity we ignore this aspect.
14. Notice that by restricting our attention to symmetric equilibrium we are excluding analysis of the dynamics of cross-section productivity along business cycles. A number of other empirical contributions focus on this point. See, for example, Caballero and Hammour (1994) and Davis and Haltiwanger (1992).
15. In the formualtion of the firm problem we have implicitly assumed that there is no 'exit effect', in other words, that the value of the firm in recession is never below τC. See Aghion and Saint-Paul (1998a, p. 326).
16. The relation between this formulation and the original approach by Romer (1986a) is discussed in Greiner and Semmler (1996a, p. 82).
17. As usual, it is assumed that the per capita utility function is such that $u'(\cdot) > 0$, $u''(\cdot) < 0$, $\lim_{c \to 0} u'(\cdot) = 0$, $u''(c)c/u'(c) \equiv -\sigma$, constant.
18. Clearly, conditions (33)–(34) are also sufficient for an optimum if the following transversality condition is satisfied

$$\lim_{t\to\infty} \exp\left[-\left(\beta-n\right)t\right]\lambda\left(t\right)\left(k(t)-k^{*}(t)\right)\geq 0$$

where k^* is the optimal value of per capita capital stock.

19. For an introduction to the concept of Hopf bifurcation, see Gandolfo (1997, ch. 25).
20. From condition (41) it follows that $\alpha + \sigma < 1$ is a necessary condition for a Hopf bifurcation to occur.
21. To obtain these conditions we have taken account of the fact that, in equilibrium, we must have $h = h_a$.
22. Benhabib and Perli (1984, p. 117–18) have shown that the maximised Hamiltonian is jointly concave in (k, h). This implies that the above conditions are also sufficient for problem (43).
23. It is possible to show that these values satisfy the transversality condition. See Benhabib and Perli (1994, p. 122).

17. Growth theory and the environment: how to include matter without making it really matter

Tommaso Luzzati

17.1. INTRODUCTION

It is sometimes useful to start by asking why questions are asked. In this respect, the question of the limits to growth is a very ‘natural’ one. Our everyday experience is that growth eventually ceases in any process, at least as concerns natural processes. At a more abstract level, this has a correspondence in sciences, where growth processes are modelled or conceived as exhibiting logistic (or similar) trends. Examples ranges from biology, e.g. cell cultures, to business analyses, e.g. the life-curve of products.

Two mechanisms that end growth in natural processes are easily identified. One is due to the lack of nutrients/inputs or other external factors feeding the process. The other one is internal, the inhibition that occurs due to (by) products of the process, a ‘poisoning’.

As regards environmental limits to economic growth we can find both types of limits. Scarcity of resources might leave production without its material basis, while waste and pollution might ‘poison’ the environment within which the economic process takes place and make further economic growth undesirable.

The classical economists perceived the limits from natural inputs – the times were not ripe enough to think of a ‘self-poisoning economy’ due to environmental degradation. What is relevant here, however, is that the question of ‘limits to growth’ was central to their research agenda. Their analyses of the increasing size (and structural changes) of their economies also included the causes that would have stopped it. That growth would have stopped was not in dispute. Ricardo, for example,

> disliked the idea of the stationary state […] and saw two factors that might, at least temporarily in his view, delay the stationary state. The first was international trade […]. The other […] was technical change […]. However, Ricardo, like

> contemporary theorists of limited economic growth, viewed both free trade and technical change as only temporary stopgaps delaying, but not preventing, the arrival of the stationary state (Foley and Michl, 1999, p. 161)

Industrial revolution – its deep specialisation, its strong urbanisation and its opulence – dissociated 'Western man' progressively from his environment and from nature in general. This affected many economists too as, from the neoclassical revolution onwards, most economic theory (see Martinez Alier, 1987 for the exceptions) disregarded the question of the material basis of economic process. After a century man looked so powerful that economists had an almost unanimous reaction when 'The Limits to Growth' (Meadows et al., 1972) appeared. They regarded such a question as not to be put on the research agenda since, consistent with their experience, they perceived unlimited economic growth as obvious. This was a U-turn with respect to classical economists.

Forced to tackle the question, mainstream economics gave its answer shortly after the publication of the report in question. The analytical framework was the well-established neoclassical growth theory while the focus, consistent with the ongoing debate, was on exhaustible resources. The new growth theory has tackled the question again since the 1990s. However, in a world where waste and pollution are leading actors, economists understood that lack of inputs can be a minor problem as compared to environmental deterioration. As a consequence, the environment is modelled mainly as a renewable resource and beliefs are more cautious, that is, the existence of environmental limits is considered at least as possible.

Does the mainstream growth theory hit its target, to show that 'there might actually be no environmental limits to economic growth'? The present paper aims to answer this question.

17.2. THE FIRST REACTION TO 'LIMITS TO GROWTH'

It is well known that the report 'Limits to Growth' provoked strong criticism adverse reactions from many economists who considered it nonsense (see, e.g., Beckerman, 1972). If they did not show methodological consistency,[1] it is also true that much criticism was forcefully argued. At the same time it was mainly based on the narrow interpretation of the report that prevailed in the general debate. In particular the focus remained on its quantitative predictions (interpreted as actual rather than *ceteris paribus* predictions) and on the problem of resource scarcity (it was about the time of the first oil crisis). Both the qualitative mechanisms outlined in the analysis and its attention towards the issue of pollution as well as the general interdependence between economic processes and ecosystems failed to receive much attention (Common, 1997).

Consequently, the problem was initially conceived and modelled by economists as being mainly one of economic growth with exhaustible resources. Such a problem fitted well within the mainstream approach, of which a major issue is the allocation of scarce resources. The answer within that framework was substitution of scarce inputs of production via changes in relative prices. Limits to growth would not be binding if, in the long run, exhaustible inputs were (markedly) substituted. In a market economy, the occurrence of such a process was seen as highly probable. Any increasing scarcity of particular materials would lead to a progressive increase in relative prices. Forced by this pressure, technology and science would find better extraction techniques, new inputs, and new productive processes. At the same time, progress in abatement techniques would reduce pollution.

Technological progress actually occurred, materials reservoirs were shown to be greater than expected, new materials and processes are being continually introduced, production is often cleaner. However, the argument still looks weak as nothing guarantees that progress will be strong and quick enough to prevent resource scarcity from becoming binding. To believe that 'salvation' will come from advancement in technology is an extrapolation of what occurred in the recent past based on *static expectations*, resembling more an act of faith in human power than a scientific argument.

Doubts in that faith also emerge when looking at the formal models built. One reason for being suspicious is the simplistic view of the environment as a mere resource input to production. On the contrary, the environment and ecological systems provide many functions for humans and life in general, in a setting where there are sizeable risks of irreversibility, thresholds and catastrophes. Thus, technological progress needs not merely to find substitutes for resources, but also to offset the general damage generated by a highly disturbed environment on human society. The analysis should then also consider the possibility of a halt due to 'poisoning'. Another reason concerns the realism of the conditions that are found to be necessary for unlimited growth. As we will see, this applies also to the analyses developed in the 1990s within the New Growth Theory.

The first strand of literature is well represented by three seminal papers, Dasgupta and Heal (1974), Solow (1974), and Stiglitz (1974), published in a special issue of the Review of Economic Studies. They develop one-sector models of neoclassical growth with exogenous technical progress where, along with capital and labour, also a depletable resource is included in the production function. Production is specialised with CES functions. Utility depends only on consumption. A central planner maximises the present value of the utility function of the representative agent.

The results depend crucially on the elasticity of substitution between capital services and the resource. For elasticity greater than one, production is possible even without the resource. The depletable resource does not constitute a limit since it can be progressively substituted by capital. For elasticity less than or equal to one these models allow sustained utility or even optimal growth only in the presence of a strong enough and increasing resource-augmenting technical progress (see Stiglitz, 1974).

Neither way to circumvent the problem seems very convincing. It is hard to think that the whole economic process could run without matter/energy. On the other hand, with Toman et al. (1996, p. 146), one can doubt

> whether it is realistic to make such a conception of technical progress that squeezes a constant flow of [...] services out of an evershrinking flow of resource service inputs.

In both cases the issue is whether it is realistic for the amount of used resources to tend asymptotically to zero, which also entails a ratio of 'exhaustible resource to Income' that tends to zero. The basic requirement for unlimited growth in this class of model is therefore a very simple one, a progressive decoupling of income from its material basis.

17.3. A BETTER REPRESENTATION OF THE ENVIRONMENT

From the 1970s, economics gradually broadened its view about the relation between the environment and the economy. First, pollution started to be included in the models (e.g. Forster, 1980). More importantly, the general interdependence between ecosystems and economies started to become accepted. A new discipline was born towards the end of the 80s, Ecological Economics, in order to take explicitly into account this interdependence. In general terms, it was acknowledged that the environment is not merely a source of resources. The environment provides services of waste absorption and general ecosystem maintenance; it directly enters the utility function both due to its amenity value and its effects on health, and it affects production. Moreover, pollution abatement is a major economic activity.

At the same time, confidence in the non-existence of environmental limits to growth started to decline. The progressive deterioration of the environment, the appearance of the idea of sustainable development (WCED, 1987) and, perhaps, the collapse of the Soviet Union, raised the issue of the possible environmental bankruptcy of the market economy. Environmental degradation was becoming increasingly evident, global, and damaging, the market economy was no longer under discussion, the notion of sustainable

development suggested the possibility of making market growth and the health of ecosystems compatible. 'No environmental limits to growth' was not warranted anymore and became rather a desired goal. In this new atmosphere it was possible for a joint group of economists and ecologists (Arrow et al., 1995) to work together and agree that there are 'limits to the carrying capacity of the planet' (Ibidem, p. 520).

17.4. NGT AND THE ENVIRONMENT

When introducing the chapter on growth and the environment in their textbook Aghion and Howitt claim that endogenous growth theory

> is also inherently more suitable for addressing the problems of sustainable development than is the neoclassical theory, because whether or not growth can be sustained is the central question to which endogenous growth theory is addressed (Aghion and Howitt, 1998, p. 151).

Nonetheless a quick look both at reference databases and at textbooks reveals that the environmental question is not so central in the research agenda of new growth theory. Most recent surveys include the above-mentioned chapter by Aghion and Howitt (1998), and Smulders (1999). These surveys contain two archetypal models used by new growth theory to take the environment into account. This section will first describe and compare them, and then attempt to provide a critical assessment.

The first framework was set forth by Stokey (1998), in a paper aimed at providing analytical foundations for the Environmental Kuznets Curve.[2] Aghion and Howitt further elaborated on it. In this framework the critical environmental factor is pollution. Pollution (treated as a flow or as a stock) negatively affects utility while being a joint product of production.[3] A growing income will then increase both consumption and pollution. Whenever, as reasonable, the welfare increase due to an increase in consumption is more than offset by the welfare loss due to an increase in pollution, then economic growth will become undesirable. The way out that Stokey suggests is a progressive shift towards cleaner technologies, whose adoption, however, is costly as it entails reductions of output.

The technology side is as follows:[4]

$$Y=f(\cdot)z \tag{1}$$

Where Y is income z, $z\in[0,1]$, is a (direct) index of pollution intensity of production, $f(.)$ is potential income. This equation reflects the assumption that cleaner technologies entail costs.

Pollution,[5] X, is assumed to be an increasing and convex function of actual output, given potential output $f(\cdot)$. Stokey uses the following

$$X = f(\cdot)z^{\beta} \qquad \text{with } \beta > 1 \tag{2}$$

By combining the two equations, actual output can be obtained as

$$Y = f(\cdot)^{1-1/\beta} X^{1/\beta} \tag{3}$$

which highlights the role of pollution as a production factor in this model.

A standard utility function, whose arguments are consumption and pollution,[6] and specific functional forms for potential output, $f(\cdot)$, complete the model. Stokey analyses two different functions for $f(\cdot)$, an AK production function, $f(\cdot) = AK$, and a Cobb–Douglas with exogenous technical progress, $f(\cdot) = AK^{\alpha}e^{gt}$ with $0<\alpha<1$.

To get an intuition of the outcome of the model it must be noted that the marginal product of capital is MPK = $B\,Y/K$, where $B=(1-1/\beta)$ in the AK case, and $B=\alpha(1-1/\beta)$ in the Cobb–Douglas one. If $f(\cdot)=AK$ then the average product of capital is Az, implying $\mathrm{MPK}=(1-1/\beta)Az$. As optimal z exhibits a decreasing path,[7] the MPK will fall below the rate of time preference and make investment not attractive anymore. Growth, while being technologically feasible (it is possible to choose a rate of change of z such that output grows and pollution declines), is not optimal. On the contrary, in the presence of exogenous technical progress the optimal outcome can be unbounded growth. This is because the average product of capital, $\mathrm{MPK}= \alpha(1-1/\beta)Y/K$, is constant along the balanced growth path. Similar conclusions are obtained within the endogenous growth Schumpeterian framework developed by Aghion and Howitt (1998, p. 151–71).

A distinctive feature of Stokey's paper is that pollution can increase without bounds. This is unrealistic due to the existence of critical ecological thresholds 'below which environmental quality cannot fall without starting in motion an irreversible and cumulative deterioration entailing a prohibitive cost' (Ibidem, p. 157). In the presence of such thresholds, income can grow without bounds only if optimal pollution intensity, z^*, tends to zero at an appropriate speed (see eqs. (1) and (2)). With a CES utility function additively separable in consumption and pollution, this is optimal only if the elasticity of the marginal utility of consumption is greater than one. The same holds for the Schumpeterian model (see Ibidem, p. 161).

Smulders' (1999) archetypal model starts by modelling the environment as a renewable resource. It assumes that it is possible to define a variable, environmental quality, N, that, as usual, follows a spontaneous logistic growth trend[8] that is altered by 'extraction' of resource, R, for production purposes.

$$\dot{N} = g(N) - R \qquad \text{with } g(0)\le 0,\ \ g_{NN} < 0 \tag{4}$$

Despite appearances, this is not very different from the Stokey/Aghion–Howitt framework. When modelling pollution as a stock, P, its dynamic will be ruled by the absorption capacity of the environment, $h(P)$, which again can be assumed to show a logistic trend, and by the rate of pollution (i.e. pollution added by new production). Thus, one can write a differential equation[9] analogous to (4)

$$\dot{P} = h(P) - X(Y,z) \qquad \text{with } h(0)\le 0,\ \ h_{NN} < 0 \tag{5}$$

The analogy goes beyond mere formality as pollution stock is in fact inversely related to environmental quality. Moreover the rate of pollution, X, is positively related to the extraction rate, R, as, due to the law of conservation of mass and the degradation of matter caused by the economic process, pollution is by the amount of material inflows (extraction).

With regard to preferences, Smulders assumes, as usual, that the representative agent has a utility function that depends positively both on consumption, and on environmental quality.[10]

Production, Y, is a positive function of extracted resources, state of the environment, and capital. All inputs are essential. The long-run marginal productivity of capital does not go to zero. As usual, production is partly consumed and partly accumulated within new capital. Capital, H, includes both physical and human capital

$$Y(N, R, H) \tag{6}$$

with $Y(0,R,H) = Y(N,0,H) = Y(N,R,0) = 0$ and $Y_N \ge 0$, $Y_R > 0$, $Y_H > 0$

Given this setting, unlimited growth with a non-deteriorating environment can be optimal (in the standard representative agent framework and depending on the utility function) if the economy lives off a constant level of 'extraction' which is consistent with ecological stability, $R=E(N)$ (see eq. (4)). Constant levels of both environmental quality and resources enter production while the economy can be optimally fuelled by levels of man-made capital, which are optimally increasing in virtue of the absence of decreasing returns. This is illustrated by a simple example Smulders (Ibidem, p. 613–14).[11] illustrates this through a simple example where constant returns to human capital are assumed. In this case the production function becomes $Y(N,R,H)=y(N,R)H$ so that the whole model ends up as an AK model. If, for simplicity's sake, a constant saving propensity is assumed, s, the accumulation equation becomes $\dot{H} = sy(N,R)H$. Dividing by H, the long-run growth rate of a balanced growth path is obtained.

This brief analysis has shown that the two frameworks described in this section differ only with respect to the 'extended' production function, i.e., the production function that also includes the 'use' of the environment. Such differences explain the different theoretical predictions.

It may be interesting to look at the mechanisms behind the two production functions. With reference to Stokey/Aghion–Howitt one can raise the doubt, as an external critique, that technology at the level of the whole economy can be considered as a choice variable. 'Production' of new technology needs time and is the outcome of an evolutionary process. When looking specifically at the model, equation (1) does not seem to have robust justifications since a cleaner technology does not necessarily entail a cost, at least in static terms. This is true both at the individual level (e.g. a single production process that is less polluting due to a reduced use of materials is actually cheaper) and in terms of aggregate output (e.g. abatement costs enter GDP, contributing positively to the relationship between cleaner technology and higher income). Secondly, as emphasised by Aghion and Howitt themselves (Ibidem, p. 162), the above-mentioned condition on the elasticity of marginal utility of consumption (>1) is problematic in many macro-economic models. Finally, as for neoclassical growth theory (see Section 17.2) with exhaustible resources, doubts can be raised about the consistency of the model with physical laws. Actually it is particularly striking that the pollution intensity for the whole economic process is assumed not to be bounded from below.

Smulders' archetypal model is neater as it allows us to go straight to the central issue for unbounded growth, the need for the economy to be run by processes whose interference with the environment becomes small to the point of disappearing. Smulders directly assumes that production does not necessarily entail pollution as a joint product. Given a minimum requirement of natural inflows, *R*, production can be increased indefinitely by increases in (human) capital, *H*, is assumed not to affect the environment.

The following conclusion can be drawn. Although endogenous growth models avoid simplistic representations of the links between the economy and the environment, the conditions for unlimited growth are built on attempts to break exactly those links, that is, on attempts to decouple matter from the economy. This is obtained in one case, Stokey and Aghion–Howitt, by assuming that technology can make production non-polluting,[12] and in the other, Smulders, by directly assuming that production need not affect the environment.

As it is undesirable to have ever-increasing pollution, and as there are limits to the cleanness of average technology, the conditions for continual growth identified by the new growth theory can be summarised by the idea that the economy must become decoupled from matter. Material inflows

must be bounded from above so that growing income will make the ratio 'Material Inflows'/'Income' fall asymptotically to zero. As we will see in the next section, this requirement (which is similar to that found in the neoclassical growth theory with depletable resources, see Section 17.2.) can be obtained in a simpler way.

17.5. BOUNDED MATERIAL INFLOWS

As Georgescu-Roegen emphasised, the economic process consists in dissipating matter and energy, that is, in producing waste. Thus, material inputs[13] matter not so much as input for production, but as they start becoming waste from the very beginning of their use.

Bearing this in mind, it is easy to understand how economic growth can occur. Let *I* be the impact, the harmfulness, of the human system. It is reasonable to impose, $I \le I^M$ as beyond I^M the impact on human welfare is undesirable and/or the ecosystems enter a catastrophic involution. z is the 'dirtiness' of the technology, bounded to be strictly positive for the reasons seen before, $z \ge z^m > 0$. *M* is total material use. As in the debate during the 1970s, let the impact be given by the product of some measures of the material scale, *M*, and the state of the technology. Then the following inequalities hold:

$$I^M \ge I = Mz \ge Mz^m \tag{7}$$

implying an upper bound to *M*, $M \le I^M/z^m$. *M* is unbounded only if one assumes there is no upper bound in *I*, or if z^m tend to zero, exactly those hypotheses used in the Stokey and Aghion–Howitt approach.

In a more descriptive way the need to decouple matter and income can be summarised as follows. Material inputs, after remaining for a while within the economy, go back to nature as waste. This affects the natural environment in ways that can be harmful to human welfare both directly and indirectly via reduced productive efficiency:

$$\xrightarrow{a} \text{INFLOWS} \xrightarrow{b} \text{OUTFLOWS(waste)} \xrightarrow{c} \text{HARMFULNESS}$$

There are three ways to reduce the negative consequences on welfare. First, one can break the link '*c*', between waste and its harmfulness by improving the 'quality' of waste, thanks to 'end-of-pipe' tools, such as treatment of toxic waste, emission abatement, and so on. Second, there is some space for increasing the persistence of matter within the economic system (arrow *b*), by promoting, for example, product durability, repairing, and recycling.

Third, less material can be used (arrow *a*) thanks, for example, to increases in material efficiency or in durability (*b*).

If material inflows/outflows are left to follow the growth of income, given the limits to the possibility of preventing goods deteriorating into waste, and the limits to the 'regenerative' capacity of nature, the improvement in 'end-of-pipe' technologies must make the average impact per unit of waste tend to zero. The lack of plausibility of an economy running on 'non-polluting' technology has already been discussed in the previous section. Consequently, if income has to grow indefinitely (and as long as our environment remains our planet), material (and energy) throughput of the economy must be bounded from above. In other words, income has to become decoupled from its material basis and the ratio material input / income has to tend to zero.

The answer to the question whether it is possible for income to grow indefinitely given a constant amount of material inflows 'lies in what we mean by output' (Petith, 2001, p. 15). Somewhat surprisingly, this question has not received much attention in the debate, including the recent forum 'Georgescu Roegen vs Solow/Stiglitz' organized by Daly on *Ecological Economics* in 1997 (issue 22). Actually physical output, income and welfare are almost used as synonyms while, on the contrary the real issue is how our society maps physical output both onto income and onto value/welfare. Prof. G. Fuà, in his last masterpiece, taught us how problematic it is to look for a mapping between income and welfare in rich countries (Fuà, 1993). To add 'matter' to the picture looks even more problematic.

However some considerations can be made on what is occurring and its perception. As regards perception it is illustrative to look again at Smulders (1999, p. 610) and read in the introduction that

> environmental and natural resource constraints did not turn the historical growth process into stagnation. Instead, accumulation of human knowledge […] allowed the economy to expand within the fixed physical system of the earth. […] (man) continually creates new knowledge to derive more value from *a given amount of physical resources* (emphasis added).

Such a position reveals firstly a sort of 'linear thinking' as it seems to suggest the prediction of perpetual growth merely based on the (relatively recent) past. More importantly, the available amount of physical resource is not distinguished by the amount of resources actually used. This is not to deny that knowledge can be considered the engine of growth. However, the fuel has been the dramatic growing amount of materials (particularly fossil energy materials[14]) occurring in the past two centuries. More value has been created from a *growing*, rather than a *given*, amount of physical resources.

However, Smulders cannot be blamed for the above statements as they constitute a shared perception,[15] the perception that the economy has started

to decouple. This feeling underlies, for example, relevant empirical research that is 'desperately seeking (environmental) Kuznets'.[16] The idea we are entering a 'weightless economy'[17] is due both to progress in pollution abatement techniques and the appearance of many new products of the 'knowledge economy', where high value is embedded in a few bits or in some lines of code.

The fact that this is probably an illusion may be easily understood by examining the literature on the EKC (see references mentioned in note 2) or by looking at the increasing trends in material requirements of the developed economies (see, e.g., Adriansee et al., 1997). The reasons for this illusion are manifold. For example, although some new products are almost immaterial, it cannot be forgotten that their consumption needs complements which are highly material. Software alone is useless, we need hardware, whose production requires a large quantity of material inflows, which soon becomes obsolete due to progress in software itself. Another reason is the small size of many end-products, which is often misleading as small size does not necessarily entail low material inflows. Generally speaking, the wrong perception seems to be arising from a mistake in shifting from one hierarchical level to another, as small improvements visible at the individual level are believed to hold also at the level of the whole economy (see, for example, Giampietro and Mayumi, 2000).

Future trends of total throughput are uncertain. There could be a reduction in the throughput of developed countries, although such a reduction would easily be more than offset by the increase in the throughput of the rest of the world as and when it starts growing.

17.6. CONCLUSION

Aghion and Howitt claim that

> … endogenous growth theory […] does imply that with enough innovations, and the right direction of innovations, such an outcome (sustainable development) is at least *within the realm of possibility* (Aghion and Howitt, 1998, p.151, emphasis added),

while Stokey (1998) is even more optimistic about environmental limits to growth. However, for the reasons seen above, the conditions that the growth theory sees as necessary for unbounded growth do not look very much 'within the realm of possibility'. This is not a novelty in economics, as both rationalisation rather than explanation is often the theoretical purpose and the realism of hypotheses is seldom an issue.

Moreover, there is a problem of methodological consistency. The principle of Occam's razor, often invoked by economists, constitutes a problem for most of the literature surveyed here. A major outcome of its elegant formal models and optimisation techniques, unbounded growth, can be obtained in a much simpler way. The obvious necessary condition for unbounded growth is to make the economy ultimately grow out of a constant material throughput. Whether or not this can occur is an open question. However, we need to be conscious that it is not true that economic growth was supported by a finite amount of physical resources. Material throughput (especially fossil fuels) has dramatically increased and is still on the increase.

NOTES

1. As was emphasised by Georgescu-Roegen (1976, p. 21–2) much of what neoclassical economists said could also have been used against their analysis.
2. The EKC is a supposed hump-shaped empirical relationship between income and pollution, a relationship that, despite claims by the author, is far from being proved (see, for example, Stern, 1997; de Bruyn and Heintz, 1999; de Bruyn, 2000).
3. Wastes could affect production so greatly as to make growth not even feasible. However, as the most immediate effects of pollution are on welfare, the focus is here (as in the literature) on the optimality of growth.
4. Notation in Stokey is slightly different.
5. Alternatively, 'increase in pollution' when pollution is modelled as a stock.
6. The utility is assumed additively separable in consumption and pollution.
7. Stokey shows that optimal *z* is decreasing in potential outcome, *f*(.).
8. As E is continuous and $E(0) \leq 0$, the existence is admitted of a threshold below which the renewable resource enters a process of progressive deterioration.
9. Actually, Aghion and Howitt use an equation similar to (4) except for using 'pollution' instead of 'extraction rate'.
10. $U(C, N)$ with $U(C,0) = U(0, N) = -\infty$ and $Uc > 0$, $UN \geq 0$
11. If constant returns to human capital are assumed the production function becomes $Y(N,R,H) = y(N,R)H$ (the whole model ends up as an AK model). With a constant saving propensity, s, the accumulation equation becomes $\dot{H} = sy(N,R)H$. Dividing by H, the long-run growth rate of a balanced growth path is obtained.
12. More precisely, in Stokey matter does not matter also in another way, as, depending on the utility, pollution can grow without bounds without consequences for ecological systems.
13. Material inputs are to be considered key indicators of the interference provoked by man on natural processes, as is particularly emphasised by research on dematerialization (see, e.g., Schmidt-Bleck, 1994) which started in the 1990s at the Wuppertal Institute for Climate and the Environment.
14. See, e.g., Cohen (1995).
15. The same confusion between available and actually used resources is also in Aghion Howitt, 1998, p. 151: 'If it had not been for resource-saving innovations it is unlikely that our *finite* planet could have supported the expansion in material welfare' (emphasis added).
16. This is the title of a recent paper by Galeotti and Lanza (1999).
17. The term 'weightless economy' is used, e.g., by Quah (1999).

18. Modelling growth and financial intermediation through information frictions: a critical survey

Salvatore Capasso

18.1. INTRODUCTION

Considerable empirical evidence has shown strong linkages between real and financial development. As economies grow, the relative size and complexity of financial systems tend to increase. New markets and financial instruments develop, while the role of financial intermediaries tends to change. Financial intermediation, very limited in the early stages of economic development, becomes increasingly important with economic growth. However, as economies continue to grow, better organised financial markets facilitate the direct transfer of resource between lenders and borrowers: stock markets develop and financial intermediaries play a decreasing role, in relative terms, in the credit market.

The idea that financial markets affect the real allocation of resources and influence capital accumulation and growth is a very old one in economics. Bagehot (1862) firmly believed that capital in England was more productive than in other countries because, in England, larger and better organised capital markets were channelling resources towards more productive investments. Schumpeter (1934), on the other hand, stressed the role of financial intermediation, and in particular of banks, in improving resource allocation and enhancing the aggregate productivity of capital. More recently, Hicks (1969), in highlighting the importance of financial markets, suggested that the industrial revolution was not the result of innovations and the development of new technologies, but rather the result of the expansion of financial systems that allowed the applications of these technologies.

In recent years, new research has attempted to provide a rigorous theoretical interpretation of the linkages between the real and financial side of the economy. In the wake of the works by Gurley and Shaw (1955, 1960, 1967), Goldsmith (1969) and McKinnon (1973), a great number of studies have attempted, in the last two decades, to give a theoretical explanation for the positive empirical correlation between financial development and growth.

In a standard Arrow–Debreu framework, in which markets are perfect, agents are fully informed and there are no transaction costs, financial markets play no role in the allocation of real resources. In order to investigate the channels of interaction between the financial and the real sector, economists have modified this framework by introducing some kinds of market frictions, such as liquidity costs, transaction costs or imperfect information. Indeed, in the presence of market frictions financial markets can affect the allocation of resources and the process of economic development. Thus, for example, financial markets can affect growth by reducing liquidity risks, or by increasing the flow of savings and by channelling such resources towards more productive alternatives (Greenwood and Jovanovich, 1990; Levine, 1991; Bencivenga and Smith, 1991; Saint-Paul, 1992; Blackburn and Hung, 1998 among others). Levine (1997) and Becsi and Wang (1997) provide an extended review of this literature.

Following the developments of the research in financial economics on the theory of the optimal financial contract, the most recent literature has focused on the assumption of information asymmetries between agents in modelling finance and growth. These studies integrate microeconomic models of optimal financial contract under information asymmetry into dynamic general equilibrium models, with very interesting results.

In the presence of information asymmetries between lenders (typically households) and borrowers (typically firms) different informational problems might arise, and the exchange of resources can become costly, sometimes to such an extent as to prevent capital markets from functioning. For example, problems of adverse selection might arise when firms have the possibility of hiding their expected profits or their level of efficiency. Problems of moral hazard, instead, might arise because of the incentive of firms to misreport the actual return on their investments. These informational problems generate agency costs, and the financial contract is the result of agents' attempts to reduce these costs. The contract, as well as the credit market structure, is, therefore, endogenously determined. The link between growth and credit market structure arises because growth can affect the level of agency costs and hence the financial arrangements, while the structure of the credit market affects growth because it determines the amount of resources invested and the allocation of capital. Thus, for instance, capital accumulation can reduce the level of credit rationing because it increases the cost of rationing. Credit rationing, on the other hand, can affect growth because it reduces the amount of savings channelled to investments.

Since these models are built on specific assumptions regarding agents' endowments and the nature of information distribution, they can display

quite different analytical structures. Moreover, such differences can be magnified by the fact that the structure of the financial market becomes significantly complex under the assumption of information asymmetries, and, therefore, specific simplifying assumptions are required.

This chapter will review the literature on finance and growth, which is fundamentally based on the assumption of asymmetric information, by critically assessing the nature and consistency of the other most common assumptions. The objective is to clarify issues related to modelling procedures, which might appear peculiar to these models, and to identify the similarities that lie behind apparently very different analytical set ups. Filling an existing gap in the literature, this chapter provides a general representation of the latest results of the research on finance and growth. Attention is focused on the elements that characterise the optimal financial contract, the structure of the real side of the economy and, most importantly, on the channels of interactions between financial and real sectors. In doing so, the chapter highlights the analytical and conceptual mechanisms of interplay between financial development and growth and the possible avenues for further research in the area. It also provides a general illustrative model which attempts to capture the basic working of these models in a single general framework.

The chapter is organised as follows. In Section 18.2 we analyse the credit market structure under information asymmetries. After classifying information asymmetries into two broad categories (*ex ante* and *ex post*), we examine, in each case in turn, the informational problems that might arise, and the possible solutions in terms of the optimal financial contract. In Section 18.3 we describe the basic analytical structure and discuss the main assumptions in these models: agents' endowments, technologies and investment opportunities. Section 18.4 studies the channels of interactions between capital accumulation and financial development and the possible dynamics of the models. In Section 18.5 we present a simple general model which attempts to compact the different features of these models in a unique analytical framework. Some concluding remarks are included in section 18.6.

18.2. THE CREDIT MARKET UNDER INFORMATION FRICTIONS

The assumption that credit markets are characterised by strong informational asymmetries between borrowers and lenders finds large support in everyday experience and has a stronghold in basic intuition. It is natural to think, for example, that an entrepreneur seeking external funding knows much more about his own activities and prospective profits than the bank or the investor that is willing to supply the funding. In addition, this assumption is not a secondary one and can have significant consequences for the agent's financial arrangements. In recent years, it has been widely shown that the optimal financial contract depends strongly on the nature of informational problems in the economy.

On logical grounds information asymmetries can be classified into two categories. In the first, lenders cannot observe the borrower's type – i.e. the borrower's characteristics – (*ex ante* information asymmetry); in the second, lenders cannot observe the outcome of the borrower's activity or action (*ex post* information asymmetry). The distinction is important given that the set of problems and the possible solutions arising in each case can be very different. Typically, *ex ante* information asymmetries generate problems of adverse selection. *Ex post* information asymmetries, instead, generate problems of moral hazard. As a consequence, the financial contracts that originate in each of these cases can substantially differ, with important implications for the interpretation of the co-evolution of financial and economic development, given that the relationship between the financial sector and the real sector depends on the nature of the financial contract.

Admittedly, even though such informational problems are completely different and easily identifiable on theoretical grounds, in practical terms they often coexist and may be very difficult to disentangle. It is only for reasons of clarity and simplification that much of the literature on finance and growth has focused on one kind of information friction at a time. We will now proceed to analyse each of these cases in turn.

18.2.1. Models that Are Based on *ex ante* Information Asymmetries

The typical framework with *ex ante* information asymmetries (Bencivenga and Smith, 1993, Bose and Cohtren 1996, 1997) involves the presence of different types of borrowers. The borrower's type is the borrower's private information and is identified with the ability of the borrower to repay the lender. Usually, this corresponds to the expected return on a project for which the borrower needs external funding. Even though the borrower's type is unobservable, a solution to the informational problem can eventually be found given that lenders know the distribution of the different types in the economy and therefore have a knowledge of the expected returns on the loan which they will eventually issue. The possible solutions, however, will also depend on other specific assumptions, such as the agent's endowments, opportunity costs, returns on technologies: assumptions that describe the general structure of the economy.

The informational gap between lenders and borrowers generates agency costs. Depending on the nature of these agency costs different contracts can arise. Under information asymmetry, a single unique contract for all borrowers (pooling equilibrium) might be unfeasible or, certainly, very costly since it risks attracting only bad or inefficient borrowers, driving the good or efficient ones out of the market (a typical lemons' problem as described by Akerlof, 1970). It has been argued, for example, in an attempt to explain credit rationing, that even in the presence of an excess demand for credit, banks might not find it optimal to increase the interest rate on their loans, but prefer to credit ration, since higher interest rates might attract a greater number of riskier agents, with the result of obtaining a lower expected return on the loans (Keeton, 1979; Stiglitz and Weiss, 1981). Certainly, it is true that under the assumption of information asymmetry borrowers might have the incentive to hide their type and pretend to be some other in order to obtain contracts with better conditions.

One possibility lenders have of solving *ex ante* informational problems is to design different contracts for each type and to make sure that each borrower will optimally prefer the contract designed for his/her own type over all the others. A logical consequence of this one-to-one matching is that the choice of the contract will be self-revealing of the borrower's type. The practical result is that borrowers are 'separated', i.e. grouped, according to their type. Of course, for separation to occur borrowers must be heterogeneous and possess specific features beyond the given differences in their types. By exploiting those differences, contracts can be made more or less attractive for one type or another. Thus, if borrowers have access to alternative investment opportunities with different rates of return and therefore have different costs when denied credit, separation can be achieved by means of credit rationing. In Bencivenga and Smith (1993), for example, and Bose and Cohtren (1996, 1997), 'bad' borrowers (with lower expected returns on their project) also have a higher alternative cost when credit rationed. As a consequence, the contract designed for 'good' borrowers can be made unattractive to bad borrowers by including a high enough probability of credit rationing.

The level of credit rationing will depend on the degree of information asymmetry, as well as on other factors, such as loan size and alternative costs. It follows that if one or more of these factors change with capital accumulation, the optimal level of credit rationing will change as well. So, for example, if the net return on the borrower's project is increasing with the loan size, the larger the loan size, the more costly it will be not to finance the project – the cost of credit rationing. Yet the loan size, the amount of resources available to be transferred, is in itself an increasing function of capital accumulation.

Separation via credit rationing is only one of the possible channels through which *ex ante* informational problems can be solved. It is realistic to think that very often lenders have the possibility to directly verify borrower's type through a costly process. Banks, for example, have all the tools to study the financial position of their clients prior to the issue of a loan. Based on this idea, Bose and Cohtren (1996, 1997) assume that lenders are endowed with a screening technology which allows the borrower's type to be determined without uncertainty and to solve, at outset and radically, the informational problem. However, screening is costly and, consequently, it is not optimal for lenders to screen in each and every contingency. The contract, therefore, will involve a probability of screening that is just sufficient to deter bad or inefficient borrowers masquerading as good or efficient ones. Crucially, in Bose and Cohtren the cost of screening depends on the loan size. This makes it clear that as the loan size increases because of capital accumulation, the return on the screening contract will change.

18.2.2. Models that Are Based on *ex post* Information Asymmetries

The situation is completely different when borrowers are homogeneous – i.e. they have the same type – but the return on their project is stochastic and unobservable; or, alternatively, when borrowers are homogeneous but it is not possible to observe their actions. In these circumstances the lender will face *ex post* informational problems.

Under *ex post* informational asymmetries, the informational gap can be filled if lenders have access to costly monitoring technology that allows them to observe the actual return on the borrowers' projects, or borrower's action. As with screening, given that monitoring is costly, it can be shown that it is not optimal for the lender to monitor in each and every state. The financial contract will therefore determine, together with all other contractual elements such as loan size and repayment, the contingency states in which monitoring occurs. Monitoring, as well as screening, is technically the means by which the contracts create the incentives for borrowers to truthfully reveal the hidden action or the actual return.

In Bernanke and Gertler (1989) the *ex post* agency cost can be lowered when borrowers can provide collateral. As in a standard *ex post* information asymmetries framework, lenders cannot observe the borrower's actual production, and hence the contract involves a probability of monitoring. The expected return on the borrower's project, the cost of monitoring and the collateral determine the set of states in which monitoring occurs. Bernanke and Gertler show that if the value of the collateral changes, as happens during business fluctuations, the probability of monitoring will change as well. Once again, this very intuitive result stresses the strict links between the real side of the economy and the financial structure.

Interestingly, it has been demonstrated (Townsend, 1979; Gale and Hellwig, 1985) that in a standard costly state verification framework (CSV), where agents are risk-neutral and monitoring costs do not depend on the project's expected return, the optimal repayment takes only the form of debt – a fixed repayment independent of the project's actual return. Thus in the standard CSV framework, apart from debt, there is no space for any other financial instruments. It is clear that, in order to explain the evolution of other forms of financial markets and instruments, this framework needs to be modified.

Following extensive empirical evidence, the most recent research has been oriented towards the description of more specific features of the linkages between financial development and economic growth. These recent studies attempt to provide an account of the development of new financial instruments that seems to be the result of capital accumulation. Economists have tried to examine how and why equity markets appear to develop relatively late in the process of capital accumulation, what happens to the debt/equity ratio when growth occurs, and the new role of financial intermediation following the development of more direct forms of lending, such as stock markets or bond markets.

Along this line of research, Boyd and Smith (1996, 1998) show that by modifying the standard CSV framework with the assumption that borrowers have access to lower return observable technology, as well as to higher return unobservable technology, the optimal financial contract will involve not only debt but also equity, a form of repayment that is a function of the project's actual return.[1] Most importantly, they show that while the issue of equity is associated to the use of observable technology, the issue of debt is associated to the use of unobservable technology. As a result, an increase in monitoring costs, which makes unobservable technology more costly, will result in a more intense use of observable technology and, ultimately, in a proportionately higher issue of equity over debt. The positive correlation between the equity/debt ratio and economic growth is explained (Boyd and Smith, 1996, 1998) under the assumption that monitoring costs depend positively on capital accumulation.

As in the case when the return on the project is unobservable, it is possible to show that debt still remains the dominant form of repayment when the borrower's action is unobservable, in which case a problem of moral hazard might arise. Very interestingly, debt might lose its dominant position if there is more than one action that cannot be observed, and the lender faces multiple moral hazard problems. Blackburn, Bose and Capasso (2001) show that under the assumption of two unobservable actions, the optimal financial contract is a combination of debt and equity. These authors also assume that one of these actions can alternatively be observed via a costly process. Therefore, lenders can either observe one action by sustaining this cost, in which case the optimal contract is a debt contract; or they can leave both actions unobservable, in which case the optimal repayment will be a combination of debt and equity. One can now understand that if this cost of observation is an increasing function of capital accumulation, the optimal contract might be only debt for low levels of capital accumulation, and a combination of equity and debt for higher levels of capital accumulation. A different explanation of the co-evolution of stock markets and economic growth thus emerges.

18.3. BASIC FRAMEWORK

Given the complex structure that credit markets can assume under information asymmetries, modelling the interrelationship between finance and economic growth requires the use of many simplifying assumptions. These assumptions are adjusted to specific modelling needs and can make the financial markets configuration widely differ from one model to another. Notwithstanding such differences, the majority of these models share common features and exhibit many similarities in their basic structure.

At the outset, one cannot but notice that in the literature the prevalent approach in modelling growth and financial development is that of ‘overlapping generations’. It is clear that when interrelationships between agents become very complex due to the presence of informational asymmetries, the overlapping generations framework greatly simplifies the understanding of such contractual arrangements and their dynamics. The results, however, do not hinge on this particular approach and could presumably be reiterated with representative agents.

Simplification is, again, the main rationale behind the common assumption of a production activity divided into two distinct processes: production of output (i.e. consumption good) and production of accumulable factor (typically physical capital). The technology describing the former process employs as inputs, among others, the physical capital produced in the latter process. This establishes the main connections between the two technologies. Most importantly, the imperfections in the credit market concern only the financing of the projects producing physical capital and not the technology producing output. As a consequence, informational problems do not concern the prices of the factors of production, such as capital and labour, as determined by the output production function. This will greatly simplify the design of the financial contract since agents take such prices as given. The rate of growth of the economy will ultimately be determined by capital accumulation as determined by the technology to produce capital.

Under a different perspective, the assumption of a double production process can be justified if one thinks of the economy as being vertically integrated, with information asymmetries concerning only the first stage of production. Two goods, capital and consumption, are produced with two different technologies, and the former is the intermediate good used in the production of the latter.

As outlined, since physical capital is used in output production, and since production of capital requires external funding, the optimal financial contract and the structure of the credit market will affect output production by affecting the flow of capital produced. At the same time, the technology for output production will affect the shape of the optimal contract, and hence the financial market structure, since it determines the prices of capital and labour and, ultimately, the equilibrium choice of the contract.

The standard working of these models is the following. Agents can be borrowers or lenders. Young borrowers have access to projects for capital production and require external funding.[2] Suppliers of such funding are young lenders who, instead, are endowed with labour. Lenders supply this labour in the sector for output production and obtain a wage income which can be lent out to borrowers. Once borrowers obtain the loan, they produce capital which is sold to output producers. It is important to mention that, while output is produced instantly (at the same time that production starts), capital takes one period to be produced and can be sold only in the period following the beginning of production. This has very important implications and may create problems of inconsistency in the results. In fact, while the price of labour (the size of the loan) is well known at the time the contract is designed, the price of capital, considered in the contract, is based on expectations.[3] It is clear that such expectations must be consistent with the actual realised return on capital. This problem can find different solutions. The most simple is to assume that the output production function displays constant returns to the aggregate capital stock, as in many endogenous growth models (Romer, 1986). In fact, if this is the case, the marginal product of capital (and the price of capital) in equilibrium will be given and constant (Bose and Cohtren, 1996; Blackburn, Bose and Capasso, 2001). Alternatively, one has to make sure that the realised price of capital is always consistent with agents’ expectations (Bose and Cohtren, 1997). As in a Nash subgame perfect equilibrium, the equilibrium optimal contract must be consistent with agents’ beliefs.

Borrowers will repay the loan to lenders once the capital has been produced. On logical grounds one can assume, indifferently, that loans are repaid in terms of capital or output. That is, the contract might bind borrowers to repay lenders directly with the capital produced, or it might require that borrowers repay lenders in units of output, after having sold the capital to output producers. The two alternatives, equivalent on logical grounds, can produce different results in practice. For instance, under the assumption that borrowers have all the bargaining power, the repayment is determined according to the lenders' reservation utility which is very often represented by an alternative investment technology. If the reservation utility is in units of output, then the price of capital will enter the repayment function only when the contract requires borrowers to repay in units of capital (and vice versa). In this case, if the price of capital is not constant in equilibrium many analytical problems might arise.

Agents are usually assumed to derive positive utility only from old age consumption. The consequence is that all first period income is invested and there is no consumption in young age. This seemingly very unrealistic assumption meets the need to eliminate problems of consumption-saving choice. Under this assumption all income is saved and, as a consequence, if there is any effect on growth from financial markets, it does not stem from savings mobilisation. This allows one to focus only on the effects of information asymmetry and credit market imperfections on capital accumulation and growth. Introducing the consumption-saving problem does not influence the marginal choice on the optimal contract, apart from the natural reduction in the availability of aggregate savings. The same rationale applies to the assumption of risk neutrality. Assuming risk-neutral agents eliminates the problem of risk sharing and allows concentration on specific informational problems.

Finally, it is worth mentioning the issue regarding bargaining power. In all these models, the contract is designed either under the assumption that borrowers have all the bargaining power (Bose and Cohtren, 1996, 1997; Bencivenga and Smith, 1993) or that lenders have all the bargaining power (Blackburn, Bose and Capasso, 2001; Boyd and Smith, 1996, 1998). Apart from empirical considerations regarding which of the two parties really has the power in credit markets, both assumptions deliver a Pareto efficient situation and are analytically equivalent. The substantial difference is that, in one case, lenders are driven to their reservation utility, in the other, borrowers.

18.4. CAPITAL ACCUMULATION AND FINANCIAL DEVELOPMENT: THE NEXUS

As outlined, under the assumption of asymmetric information in the credit market, the optimal financial arrangement is the result of agents' efforts to reduce agency costs. Given that such financial arrangements depend on specific market conditions, the optimal financial contract, and the

corresponding credit market structure, will change when those market conditions change because of economic growth. In turn, the credit market can influence economic growth since it can affect the flow of resources devoted to investment. This explains the co-evolution of real and financial development.

In every period, given the set of available information and the new market conditions, financial contracts are redesigned. Each generation of young agents (borrowers and lenders) meets in the credit market where funds are transferred from lenders to borrowers under a pre-specified contract.

The contract, as seen, depends on variables such as the expected price of capital, the amount of available savings – which, in turn, are functions of the wage rate –, and the verification costs. Accordingly, agents will design the financial contract in order to achieve the highest possible payoff and hence the highest expected utility. Of course, optimisation implies that if different forms of contract are available, the equilibrium contract will be that associated with the highest payoff. Thus, if there were two forms of available contracts to choose from, credit rationing and screening for instance (Bose and Cohtren, 1996, 1997), agents will abide, in equilibrium, by the one that delivers the highest payoff: either rationing or screening. Obviously, the equilibrium contract can also be a combination of different contracts when this combination delivers an expected payoff higher than that achievable through a single contract. If we think of equity and debt, for example, as two different forms of repayment each relating to a specific contract, a combination of the two – a repayment which consists partly of equity and partly of debt – might be strictly preferred to a repayment only in the form of debt or equity (Boyd and Smith, 1996, 1998 ; Blackburn, Bose and Capasso, 2001).

It is intuitive that the optimal contract will not necessarily be the same as capital accumulates and market conditions vary. With growth occurring, the wage rate, the price of capital and the verification costs might change, with the result that the optimal structure of the financial contracts and the associated payoffs will change as well.

The level of economic activity can be measured in different ways. As capital accumulates, given a fixed labour supply, the wage rate increases.[4] Hence the amount of resources available to lenders increases, as do the loan size and the project size.[5] The price of capital, on the other hand, decreases with capital accumulation or remains constant (in equilibrium), depending on the form of the output production function. Thus, for example, if the cost of rationing is a decreasing function of the level of economic activity (Bencivenga and Smith, 1993; Bernanke and Gertler, 1989), with growth occurring, the optimal contract might involve a lower level of credit rationing. Then again, if screening involves a cost which is a decreasing function of the loan size, agents might prefer a rationing contract to a screening one at a low level of capital accumulation, and screening to rationing at a higher level (Bose and Cohtren, 1996, 1997). This could explain why in richer economies the level of credit rationing is clearly lower than in poorer ones.

Even if monitoring costs are assumed to be fixed in units of output, they might appear to be increasing if expressed in units of capital when the price of capital is decreasing because of capital accumulation. As a consequence, with economic growth, agents prefer to reduce monitoring costs by reallocating investment from costly unobservable technologies towards observable ones (Boyd and Smith, 1996, 1998). The financial implication, as shown by Boyd and Smith, is a contract with a higher proportion of repayment in the form of equity rather than debt. Similarly, if monitoring costs are an increasing function of the wage rate, then with growth occurring, lenders might find it optimal not to sustain these costs and directly observe the borrower's actions. Hence lenders will face more complex forms of moral hazard and the financial outcome will be a contract that involves forms of financial instruments, such as equity, other than debt (Blackburn, Bose and Capasso, 2001). This mechanism could explain the emergence of stock markets at later stages of economic development.

Recalling that, usually, the output production function is separated from the capital production function, the rate of economic growth is a function of the rate of capital accumulation as determined by the capital production function. Indeed, the amount of resources invested in the latter (the amount of available savings) and, therefore, the amount of future capital, depends on the actual amount of capital available in the output production technology.[6] It is worth stressing, at this stage, that constant returns to aggregate capital in the output production function, as in standard endogenous growth models, do not necessarily imply an endogenous rate of growth. If the capital production function displays decreasing returns to capital, the rate of growth will be exogenous as in the Solow–Swan model.

As already outlined, the relationship between the level of economic activity and financial development is bi-directional. Financial markets can affect the level of economic activity by determining the amount of funds devoted to investment and capital accumulation. It is immediately obvious that credit rationing, by reducing the amount of projects undertaken or the scale of the projects, can limit capital accumulation. However, also screening and monitoring costs can lower the amount of resources accumulated when these costs consist of a net loss for the economic system and do not result in a simple transfer of resources. The general idea is that, under information asymmetries, agency costs influence investment allocation by deviating the flow of investment towards less productive activity, thereby affecting the rate

of growth. Importantly, the effect of financial markets on growth does not necessarily feed through to a permanent increase in the growth rate (Bencivenga and Smith, 1993; Bose and Cohtren, 1996). Indeed, such effects could not be analysed in a model that displays exogenous growth. Very often, as reported in the literature, the effects of financial market development on growth prove to be effects on the capital accumulation path. Thus a change in the financial contract from rationing to screening (Bose and Cohtren 1997), or from debt to debt and equity (Blackburn, Bose and Capasso, 2001) will free resources and will push the capital accumulation path upwards, temporarily boosting the rate of growth.

18.5. A SIMPLE GENERAL MODEL

We will now illustrate through a simple framework the central structure of the models on finance and growth which are based on the assumption of information asymmetries. Given the complexity of the financial market in itself and the many different assumptions made in these models, the attempt requires a large dose of generalisation. The main objective is to clarify the general working of these models and to highlight their dynamics.

The model is characterised by the following equations:

$$C = C(R_t; l_t; \pi_t; \phi_t; \gamma_t; \zeta_t) \tag{1}$$

$$R_t = E_t + D_t \tag{2}$$

$$l_t = l(w_t) \tag{3}$$

$$\pi_t = (1-\alpha)\bar{z} + \alpha z(\rho_{t+1}; w_t) \tag{4}$$

$$\phi_t = (1-\beta)\bar{g} + \beta g(\rho_{t+1}; w_t) \tag{5}$$

$$\gamma_t = \gamma(\rho_{t+1}, w_t) \tag{6}$$

$$E(K_{t+1}) = G\left[C(\cdot); \tilde{T}\right] \tag{7}$$

$$Y_t = F(K_t, L_t) \tag{8}$$

$$w_t = F_L \tag{9}$$

Equation (1) captures the form and elements of the contract, $C(\cdot)$. As already argued, the elements that a contract eventually determines include: the repayment, R_t, the size of the loan, l_t, the probability of rationing, π_t, the probability of screening, ϕ_t, and the required amount of collateral, γ_t. It is clear that while the repayment and loan size are essential elements of a credit contract (i.e. each and every contract needs to determine repayment and loan size), the presence of other contractual arrangements depends on the assumptions regarding the credit market. As seen above, under the assumption of information asymmetries, the contractual arrangement is the result of the level of agency costs, endowments and technologies, which can take very different forms. Some of these contractual elements have been explicitly considered (probability of rationing, screening etc.). However, many others could be included depending on the assumptions. The variable ζ_t captures any other element that might enter the contract.

The form of repayment is specified by equation (2). The repayment can be in the form of debt, D_t, equity, E_t, or a combination of the two.[7] Since the resources in the credit market are supplied by lenders who sell their labour to output producers, the loan size depends on the wage rate, w_t, (eq. (3)).

The probability of credit rationing, eq.(4), and the probability of screening, eq.(5), can be either constant ($\alpha = 0$; $\beta = 0$) or they can depend on the price of capital, ρ_{t+1}, and/or on the price of labour ($\alpha = 1$; $\beta = 1$). In Bose and Cohtren (1996, 1997) both the probability of screening and the probability of credit rationing are constant. In this case, the effect of growth on the credit market feeds through to a switch in the equilibrium contract (from rationing to screening), rather than the level of credit rationing (as, instead, occurs in Bencivenga and Smith, 1993). The collateral, γ_t, is a function of the price of capital and/or the wage rate. A change in either of these two variables might affect the value of the collateral and, consequently, the borrower's creditworthiness and the level of agency costs.

The expected level of capital is determined by equation (7). This depends not only on the technology to produce capital, $\tilde{T}$, but also on the contract, $C(\cdot)$. Clearly, the form of the financial contract can affect the amount of capital produced by simply determining the flow of resources devoted to investment. The technology to produce capital, on the other hand, delivers a stochastic return, which can be either observable or unobservable, and it directly affects the amount of capital produced.

The last three equations describe the output production function and the price of the factors of production, labour and capital. The output production function, as outlined, may display decreasing returns in each factor (as in a standard neoclassical growth model) or constant returns to aggregate capital (production function à la Romer). Therefore, the one in (8) is a very general production function. Equations (9) can be interpreted in the following way: if markets for capital and labour are competitive, then factors will be paid their marginal productivity.

The working of the model is the following. Given an initial level of capital stock, K_0, and the initial labour supply from young lenders, output is

produced according to (8). The output is then used to pay labour and capital their price, (9). Young lenders and borrowers meet in the credit market where, for a given wage rate and the expected price of capital, the optimal contract is determined according to (1). Borrowers who are granted credit, $l(w_t)$, produce capital $E(K_{t+1})$ which will be available in the next period. Once capital has been produced, borrowers will repay the loan. The loan will be repaid either in terms of capital or output. In the first case after producing capital, borrowers immediately repay the lenders and sell the rest of capital to new output producers. In the second case, borrowers first sell the capital to output producers and then repay the lenders. With the new stock of capital and with the labour supplied by the new young generation of lenders output is produced and the cycle restarts.

18.6. CONCLUSIONS

Empirical evidence strongly supports the view that financial markets play a key role in the process of capital accumulation. The development of financial markets seems to be positively correlated with the process of economic growth. This view is not a new one in economics. Classical economists were convinced of the importance of the financial structure for its effects on the allocation of real resources and, therefore, on the average return on investments. However, only recently have economists attempted to provide a theoretical explanation of the linkages between financial and real sectors of an economy within fully articulated general equilibrium models of growth.

The most recent research has focused attention on the assumption of asymmetric information in the credit market. Indeed, the agency costs that stem from such informational problems can determine financial arrangements that prove very different from those based on the assumption of fully or symmetrically informed agents. Most importantly, under the assumption of asymmetric information, the structure of the financial system is the result of agents' attempts to reduce these agency costs, and can be viewed as endogenously determined. As a result, this framework highlights issues that cannot be investigated and analysed in a framework where the structure of the financial market is exogenously imposed.

This chapter has provided a critical analysis of the most important assumptions in this developing literature, in an attempt to reveal the underlying common analytical framework of seemingly very different models. The inductive process which lies behind this attempt, illustrated the general theoretical mechanism that explains the interactions between financial development and economic growth.

It is important to note that the role played by financial markets, as it appears in the literature, is essentially not very far from the role that classical economists thought financial markets were playing in the economic system. Financial markets can affect the process of capital accumulation through specialisation and by channelling resources towards more productive investments. However, it is also clear that these models, so very rich in their analytical structure, have the power to highlight specific features of financial systems and, consequently, provide the possibility to understand specific aspects of the co-evolution of financial development and economic growth.

Indeed, on the basis of the potentially high explanatory power of this analytical framework, the focus of research has gradually shifted from very general questions, such as the role of financial intermediation in the process of economic development, to more specific issues, such as the endogenous development of other forms of financial instruments and markets, and their impact on growth.

NOTES

1. We recall that in the standard CSV framework the borrower has access to a 'single' unobservable technology whose return can be observed by the lender only by sustaining a monitoring, or verification, cost (for this reason 'costly state verification').
2. From now on, for simplicity, we will base the discussion on the assumption of two-period-lived overlapping generations.
3. Recall that funds to young borrowers are supplied by young lenders out of their wage income, which is well known at the time of the contract. On the other hand, the return on the project – the amount of capital produced – will only be available one period after the contract is designed.
4. The assumption of a fixed labour supply is common. The results would be the same if labour supply was growing at a lower rate than capital.
5. If there is no consumption-saving problem and all income is saved, the increase in wage results in an equivalent increase in loan size. This, of course, requires that lenders have no alternative investment opportunities with higher return and that the return on the project is a positive function of the resources invested.
6. Recall that capital takes more than one period to be produced. The amount of today's capital used in output production together with labour, determines the amount of today's savings available for investment. These resources employed in the capital production function will determine tomorrow's capital.
7. The form of repayment is assumed to be exogenously given. However, since our objective goes beyond the study of the financial arrangement in itself, and it consists in showing the relationship between real and financial development in a general equilibrium framework, these forms of repayment can be justified by the simple observation that in almost every financial system equity and debt are the most common forms of financial instruments. Moreover, there is a vast micro literature in finance that has rigorously proved that the optimal repayment takes generally either the form of debt or equity.

References

Abel, A.B. (1990), ''Consumption and investment', in B.M. Friedman and F.H. Hahn (eds), *Handbook of Monetary Economics*, vol. II, New York: Elsevier Science, pp. 725–78.

Abraham-Frois, G. (1991), 'Corporate behaviour, valuation ratio and macroeconomic analysis', in Nell and Semmler (1991), pp. 190–204.

Acemoglu, D. (1995), 'Reward structures and the allocation of talents', *European Economic Review*, **39**, 17–33.

Acemoglu, D. (1999), 'Changes in unemployment and wage inequality: an alternative theory and some evidence', *American Economic Review*, **89**, 1259–78.

Ackley, G. (1978), *Macroeconomics: Theory and Policy*, London: Macmillan.

Adriansee, A., S. Bringezu, A. Hammond, Y. Moriguchi, E. Rodenberg, D, Rogich and H. Schützt (1997), *Resource Flows: The Material Basis of Industrial Economies*, Washington, DC: World Resource Institute.

Aghion, P., E. Caroli and C. Garcia Penalosa (1999), 'Inequality and economic growth: the perspective of the new growth theories', *Journal of Economic Literature*, **37**, 1615–60.

Aghion, P. and P. Howitt (1992), 'A model of growth through creative destruction', *Econometrica*, **60**, 323–51.

Aghion, P. and P. Howitt (1994), 'Endogenous technical change: the Schumpeterian perspective', in L.L. Pasinetti and R.M. Solow (eds) *Economic Growth and the Structure of Long-Term Development*, London: Macmillan, pp. 118–31.

Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory*, Cambridge, Massachusetts: MIT Press.

Aghion, P. and G. Saint-Paul (1998a), 'Uncovering some casual relationships between productivity growth and the structure of economic fluctuations: a tentative survey', *Labour*, **12**(2), 279–303.

Aghion, P. and G. Saint-Paul (1998b), 'Virtues of bad times: interaction between productivity growth and economic fluctuations', *Macroeconomic Dynamics*, **2**(3), 322–44.

Agliardi, E. (1998), *Positive Feedbacks Economies*, London: Macmillan.

Alchian, A. (1955), ‘The rate of interest, Fisher’s rate of return over costs and Keynes’s internal rate of return’, *American Economic Review*, **45**(5), 938–45.

Allen, R.G.D. (1967), *Macro-Economic Theory. A Mathematical Treatment*, London: Macmillan.

Amadeo, E.J. (1986a) ‘The role of capacity utilization in long-period analysis’, *Political Economy*, **2**(2), 147–85.

Amadeo, E.J. (1986b), ‘Notes on capacity utilization, distribution and accumulation’, *Contributions to Political Economy*, **5**, 83–94.

Amadeo, E.J. (1987), ‘Expectations in a steady-state model of capacity utilization’, *Political Economy*, **3**(1), 75–89.

Andersen, E.S. (2001), ‘Toward a Multiactivity Generalisation of the Nelson–Winter model’, Paper presented at the 2001 Nelson and Winter Conference, Aalborg.

Anderson, G.M., C.K. Rowly and R.D. Tollison (1988), ‘Rent seeking and the restriction of human exchange’, *Journal of Legal Studies*, **27**, 83–100.

Arndt, H.Z. (1984), *Economic Development*, Chicago: The University of Chicago Press.

Arnold, L.G. (2000a), ‘Stability of the market equilibrium in Romer’s model of endogenous technical change: a complete characterization’, *Journal of Macroeconomics*, **22**(1), 69–84.

Arnold, L.G. (2000b), ‘Endogenous technological change: a note on stability’, *Economic Theory*, **16**, 219–26.

Arrow, K.J. (1962), ‘The economic implications of learning by doing’, *Review of Economic Studies*, **29**, 155–73.

Arrow, K.J. (1998), ‘Innovation and increasing returns to scale’, in K.J. Arrow, Y-K. Ng, and X. Yang (eds), *Increasing Returns and Economic Analysis*, London: Macmillan, pp. 403–8.

Arrow, K.J., B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C.S. Holling, B-O. Jansson, S. Levin, K-G. Maler, C. Perrings and D. Pimentel (1995), ‘Economic growth, carrying capacity and the environment’, *Science*, **268**, 520–21.

Arrow, K.J. and Kurz, M. (1970), *Public Investment, the Rate of Return, and Fiscal Policy*, Baltimore: Resources for the Future.

Arthur, W.B. (1987), ‘Self-reinforcing mechanisms in economics’, in P.W. Anderson, K.J. Arrow and D. Pines (eds), *The Economy as an Evolving Complex System*, Addison Wesley Publishing Company, 9–31.

Asada, T. (1995), ‘Kaldorian dynamics in an open economy’, *Journal of Economics*, **62**(3), 239–69.

Asada, T., W. Semmler and A.J. Novak (1998), ‘Endogenous growth and the balanced growth equilibrium’, *Research in Economics*, **52**(2), 189–212.

Asimakopulos, A. (1985), ‘Harrod on Harrod: the evolution of “A Line of Steady Growth”‘, *History of Political Economy*, **17**(4), 619–35.

Asimakopulos, A. and J.C. Weldon (1965), ‘A synoptic view of some simple models of growth’, *Canadian Journal of Economics and Political Science*, **31**(1), 52–79.

Atkinson, A.B. and J.E. Stiglitz (1969), ‘A new view of technological change’, *Economic Journal*, **69**, 573–78.

Atsumi, H. (1969), ‘The efficient capital programme for a maintainable utility level’, *Review of Economic Studies*, **36**, 263–87.

Azariadis, C. and A. Drazen (1991), ‘Threshold externalities and economic development’, *Quarterly Journal of Economics*, **105**, 501–26.

Backhouse, R. (1998), *Explorations in Economic Methodology*, London: Routledge.

Bagehot, W. (1962), *Lombard Street*, 1st edn 1873, Homewood IL: Richard D. Irwin.

Bairam, E.I. (1993), ‘Income elasticities of exports and imports: a re-examination of the empirical evidence’, *Applied Economics*, **25**(1), 71–4.

Bairoch, P. (1973), ‘Agriculture and the industrial revolution’, in C.M. Cipolla (ed.), *The Fontana Economic History of Europe*, vol. 3, The Industrial Revolution, Glasgow: Collins.

Baldwin, R.E. (1956), ‘Patterns of development in newly settled regions’, *The Manchester School*, **24**, 161–79.

Ball, L. (1999), ‘Aggregate demand and long-term unemployment’, *Brookings Papers on Economic Activity*, n. 2, 189–251.

Ball, R.J., T. Burns and J.S.E. Laury (1977), ‘The role of exchange rate changes in balance of payments adjustment: the U.K. Case’, *Economic Journal*, **87**, 1–29.

Baranzini, M. (1991), *A Theory of Wealth Distribution and Accumulation*, Oxford: Clarendon Press.

Barbosa-Filho, N.H. (2000), ‘A note on the theory of demand-led growth’, *Contributions to Political Economy*, **19**, 19–32.

Barreto, R. (2000), ‘Endogenous corruption in a neoclassical growth model’, *European Economic Review*, **44**, 35–60.

Barro, R.J. (1990), ‘Government Spending in a Simple Model of Endogenous Growth’, *Journal of Political Economy*, **98**(5), 103–25.

Barro R.J. and X. Sala-i-Martin (1995), *Economic Growth*, New York: McGraw-Hill.

Baumol, W.J. (1990), ‘Entrepreneurship: productive, unproductive and destructive’, *Journal of Political Economy*, **101**, 274–98.

Baumol, W.J. (1990), ‘Entrepreneurship: productive, unproductive and destructive’, *Journal of Political Economy*, **98**, 893–921.

Bean, C. (1990), ‘Endogenous growth and the procyclical behaviour of productivity’, *European Economic Review*, **34**(2–3), 355–63.

Becker, G. (1988), ‘Family economics and macro behaviour’, *American Economic Review*, **7**, 1–78.

Becker, G.S. and R.J. Barro (1988), ‘A Reformulation of the economic theory of fertility’, *Quarterly Journal of Economics*, **103**(1), 1–25.

Becker, G.S., K.M. Murphy and R. Tamura (1990), ‘Human capital, fertility, and economic growth, *Journal of Political Economy*, **98**(5), S12–S37.

Beckerman, W.B. (1972), ‘Economic development and the environment: a false dilemma’, *International Conciliation*, **586**, 57–71, reprinted in W.B. Beckerman (1992), *Growth, the Environment and the Distribution of Incomes: Essays by a Sceptical Optimist*, Aldershot: Edward Elgar.

Becsi, Z. and P. Wang (1997), ‘Financial development and growth’, *Economic Review*, Federal Reserve Bank of Atlanta, 46–62.

Benabou, R. (1996), ‘Heterogeneity, stratification, and growth: macroeconomic implications of community structure and education finance’, *American Economic Review*, **86**, 584–609.

Bencivenga, V.R. and B. Smith (1991), ‘Financial intermediation and endogenous growth’, *Review of Economic Studies*, **58**, 195–209.

Bencivenga, V.R. and B. Smith (1993), ‘Some consequences of credit rationing in an endogenous growth model’, *Journal of Economic Dynamics and Control*, **17**, 97–122.

Benhabib, J. and K. Nishimura (1979), ‘The Hopf bifurcation and stability of closed orbits in multisector models of optimal economic growth’, *Journal of Economic Theory*, **21**(3), 421–44.

Benhabib J. and R. Perli (1994), ‘Uniqueness and indeterminacy: on the dynamics of endogenous growth’, *Journal of Economic Theory*, **63**(1), 113–42.

Benhabib J., R. Perli and D. Xie (1994), ‘Monopolistic competition, indeterminacy, and growth’, *Ricerche Economiche*, **48**(4), 279–98.

Ben-Porath, Y. (1980), ‘The F-connection: families, friends, and firms and the organization of exchange’, *Population Development Review*, **6**, 1–30.

Bernanke, B. and M. Gertler (1989), ‘Agency costs, net worth, and business fluctuation’, *American Economic Review*, **79**, 14–31.

Bertola, G. (1993), ‘Factor shares and savings in endogenous growth’, *American Economic Review*, **83**, 1184–98.

Bertola, G. (1994), ‘Theories of saving and economic growth’, *Ricerche Economiche*, **48**, 257–77.

Bertola, G. (2000), ‘Macroeconomics of distribution and growth’, in B.A. Atkinson and F. Bourguison (eds), *Handbook of Income Distribution*, Amsterdam: Elsevier Science, pp. 477–540.

Besomi, D. (1999), *The Making of Harrod’s Dynamics*, London: Macmillan.

Bhaduri, A. and S. Marglin (1990), ‘Unemployment and the real wage: the economic basis for contesting political ideologies’, *Cambridge Journal of Economics*, **14**(4), 375–93.

Bhagwati, J.N. and T.N. Srinivasan (1980), ‘Revenue seeking: a generalization of the theory of tariffs’, *Journal of Political Economy*, **88**, 1069–87.

Blackburn, K. and V. Hung (1998), ‘A theory of growth, financial development and trade’, *Economica*, **65**, 107–24.

Blackburn, K., N. Bose, and S. Capasso (2001), ‘Financial development, financial choice and economic growth’, *CGBC Discussion Paper*, University of Manchester.

Blaug, M. (1975), ‘Kuhn versus Lakatos, or paradigms versus research programmes in the history of economics’, *History of Political Economy*, **7**, 399–433.

Blaug, M. (1980), *The Methodology of Economics*, Cambridge: Cambridge University Press.

Boldrin, M. and A. Rustichini (1994), ‘Growth and indeterminacy in dynamic models with externalities’, *Econometrica*, **62**(2), 323–42.

Boldrin, M. and D.K. Levine (2002), ‘Factor saving innovation’, *Journal of Economic Theory*, forthcoming.

Bose, N. and R. Cothren (1996), ‘Equilibrium loan contracts and endogenous growth in the presence of asymmetric information’, *Journal of Monetary Economics*, **38**, 363–76.

Bose, N. and R. Cothren (1997), ‘Asymmetric information and loan contracts in a neoclassical growth model’, *Journal of Money, Credit, and Banking*, **29**, 423–39.

Bose, S. (1968), ‘Optimal growth and investment allocation’, *Review of Economic Studies*, **35**, 465–80.

Bowley, M.(1975), ‘Some aspects of the treatment of capital in the wealth of nations’, in Skinner and Wilson (1975), pp. 361–76.

Boyd, J. and B. Smith (1996), ‘The coevolution of real and financial sectors in the growth process’, *The World Bank Economic Review*, **10**(2), 371–96.

Boyd, J. and B. Smith (1998), ‘The evolution of debt and equity markets in economic development’, *Economic Theory*, **12**, 519–60.

Brandolini, A. (1995) ‘In search of a stylised fact: do real wages exhibit a consistent pattern of cyclical variability?’, *Journal of Economic Surveys*, **9**(2), 103–63.

Brewer, A. (1998), ‘Luxury and economic development: David Hume and Adam Smith’, *Scottish Journal of Political Economy*, **45**, 78–98.

Britto, R. (1974), ‘Some recent developments in the theory of economic growth: an interpretation’, *Journal of Economic Literature*, **12**, 1343–67.

Brock, W.A. and S. Durlauf (2001), ‘Discrete choice with social interactions’, *Review of Economic Studies*, **68**, 235–60.

Brooks, M.A. and B.J. Heijdra (1989), ‘An exploration on rent seeking’, *The Economic Record*, **65**, 32–50.

Buchanan, J.M.. (1980), ‘Rent-seeking and profit-seeking’, in J.M. Buchanan, R.D. Tollison and G. Tullock (eds), *Towards a Theory of the Rent-Seeking Society*, College Station, Texas A&M University Press, pp. 3–15.

Buchanan, J.M. and Y.J. Yoon (2000), ‘A Smithian perspective on increasing returns’, *Journal of the History of Economic Thought*, **22**, 43–8.

Burgstaller, A. (1994), *Property and Prices: Towards a Unified Theory of Value*, Cambridge: Cambridge University Press.

Burmaister, E. and A.R. Dobell (1970), *Mathematical Theories of Economic Growth*, New York: Macmillan.

Caballé, J. and M.S. Santos (1993), ‘On endogenous growth with physical and human capital’, *Journal of Political Economy*, **101**(6), 1042–67.

Caballero, R. and M. Hammour (1994), ‘The cleansing effect of recessions’, *American Economic Review*, **84**(5), 1350–68.

Caballero, R.J. and A.B. Jaffe (1993), ‘How high are the giant’s shoulders: an assessment of knowledge spillovers and creative destruction in a model of economic growth’, *NBER. Macroeconomics Annual*, 15–74.

Canova F. (1998), ‘Detrending and business cycles facts’, *Journal of Monetary Economics*, **41**(3), 475–512.

Cantwell, J.A. (1998), ‘Innovation as the principal source of growth in the global economy’, in D. Archibugi, J. Howells and J. Michie (eds) Innovation Policy in a Global Economy, Cambridge: Cambridge University Press, pp. 225–41.

Carillo, M.R. and A. Zazzaro (2001), ‘Innovation, rent-seeking and social prestige: towards a dynamic theory of professions’, paper presented at the Conference *Old and New Growth Theories: an Assessment*, held in Pisa, October, 2001.

Cass, D. (1965), ‘Optimum growth in an aggregative model of capital accumulation’, *Review of Economic Studies*, **32**, 233–40.

Cass, D. and K. Shell (1976), ‘The structure and stability of competitive dynamical systems’, *Journal of Economic Theory*, **12**, 10–18.

Cassel, G. (1903), *The Nature and Necessity of Interest*, London: Macmillan.

Cassel, G. (1932), *The Theory of Social Economy* (first German edn 1918), New York: Harcourt Brace.

Casson, M. (1993), ‘Cultural determinants of economic performance’, *Journal of Comparative Economics*, **17**, 418–42.

Champernowne, D.G. (1945–56), ‘A note on J. v. Neumann’s Article “A model of economic equilibrium”‘, *Review of Economic Studies*, **13**, 1–9.

Charasoff, G. (1910), *Das System des Marxismus: Darstellung und Kritik*, Berlin: H. Bondy.

Chesnais, J. (1992), *The Demographic Transition*, Oxford: Oxford University Press.

Chick, V. (1995), ‘Is there a case for post Keynesian economics?’, *Scottish Journal of Political Economy*, **42**(1), 20–36.

Chirinko, R.S. (1993), ‘Business fixed investment spending: modeling strategies, empirical results, and policy implications’, *Journal of Economic Literature*, **31**(4), 1875–911.

Christiano, L. and M. Eichenbaum (1992), ‘Current real-business-cycle theories and aggregate labour-market fluctuations’, *American Economic Review*, **82**(3), 430–50.

Chu, C.C.Y. (1998), *Population Dynamics. A New Economic Approach*, Oxford: Oxford University Press.

Ciampalini, A. and F. Vianello, (2000), ‘Concorrenza, accumulazione del capitale e saggio del profitto’, in M. Pivetti (ed.), *Piero Sraffa*, Roma: Carocci, pp. 363–98.

Ciccone, R. (1986), ‘Accumulation and capital utilization: some critical considerations on Joan Robinson’s theory of distribution’, *Political Economy*, **2**(1), 17–36.

Ciccone, R. (1987), ‘Accumulation, capacity utilization and distribution: a reply’, *Political Economy*, **3**(1), 97–111.

Cigno, A. (1981), ‘Growth with exhaustible resources and endogenous population’, *Review of Economic Studies*, **48**, 281–87.

Cigno, A. (1984a), ‘Consumption vs. procreation in economic growth’, in G. Steinmann (ed.), *Economic Consequences Change of Population in Industrialised Countries*, Berlin, Heidelberg, New York: Springer-Verlag, pp. 29–49.

Cigno, A. (1984b), ‘Further implication of learning-by-doing: the effect of population on per-capita income’, *Bulletin of Economic Research*, **36**(2), 97–108.

Cohen, J.E. (1995), *How Many People Can the Earth Support?*, New York and London: Norton.

Colander, D.C. (ed.)(1984), *Neoclassical Political Economy*, Cambridge: Ballinger Publishing Company.

Cole, L., G. Mailath and A. Postlewaite (1992), ‘Social norms, saving behaviour and growth’, *Journal Political Economy*, **104**, 946–59.

Coleman, J. (1988), ‘Social capital in the creation of human capital’, *American Journal of Sociology*, **94**, S95–S121.

Commendatore, P. (1994), ‘Sulla esistenza di un’economia a due classi in un modello Post Keynesiano di crescita e distribuzione con settore pubblico e attività finanziarie’, *Studi Economici*, **51**(3), 5–38.

Commendatore, P. (1999a), *Government Deficits (and Financial Activities) in Post Keynesian Theories of Growth and Distribution*, PhD dissertation, University of Naples ‘Federico II’.

Commendatore, P. (1999b), ‘Pasinetti and dual equilibria in a post Keynesian model of growth and institutional distribution’, *Oxford Economic Papers*, **51**(1), 223–36.

Committeri, M. (1986), ‘Some comments on recent contributions on capital accumulation, income distribution and capacity utilization’, *Political Economy*, **2**(2), 161–86.

Committeri, M. (1987), ‘Capital utilization, distribution and accumulation: a rejoinder to Amadeo’, *Political Economy*, **3**(1), 91–5.

Common, M.S. (1997), ‘Roles for ecology in ecological economics and sustainable development’, in N. Klomp and I. Lunt (eds) *Frontiers in Ecology: Building the Links*, Amsterdam: Elsevier Science.

Cooley, T.F. and E.C. Prescott (1995), ‘Economic growth and business cycles, in T.F. Cooley and E.C. Prescott (eds), *Frontiers of Business Cycles*, New Jersey: Princeton University Press, pp. 1–38.

Corneo, G. and O. Jeanne (1997), ‘On the relative wealth effects and the optimality of growth’, *Economic Letters*, **70**, 37–51.

Cornwall, J. and W. Cornwall (1994), ‘Growth theory and economic structure’, *Economica*, **61**, 237–51.

Costabile, L. and B. Rowthorn (1985), ‘Malthus’s theory of wages and growth’, *Economic Journal*, **95**, 418–37.

Cozzi, G. (1998), ‘Culture as a bubble’, *Journal of Political Economy*, **106**, 376–94.

Cugno, F. and L. Montrucchio (1982), ‘Cyclical growth and inflation: a qualitative approach to Goodwin’s model with money prices’, *Economic Notes*, n. 3, 93–107.

Currie, L. (1997), ‘Implications of an endogenous theory of growth in allyn young’s macroeconomic concept of increasing returns’, *History of Political Economy*, **29**, 413–43.

Dahan, M. and D. Tsiddon (1998), ‘Demographic transition, income distribution, and economic growth’, *Journal of Economic Growth*, **3**, 29–52.

Daly, H.E. (ed.) (1973), *Economics, Ecology, Ethics: Toward a Steady State Economy*, San Francisco: Freeman.

Dalziel, P.C. (1989), ‘Cambridge (U.K.) versus Cambridge (Mass.): a Keynesian solution of “Pasinetti’s paradox”: comment’. *Journal of Post Keynesian Economics*, **11**(4), 648–53.

Dalziel, P.C. (1991–92), ‘Does government activity invalidate the Cambridge theorem of the rate of profit? A reconciliation’, *Journal of Post Keynesian Economics*, **14**(2), 225–31.

Dalziel, P.C. (1991a), ‘A generalisation and simplification of the Cambridge theorem with budget deficits’, *Cambridge Journal of Economics*, **15**(3), 287–300.

Dalziel, P.C. (1991b), ‘A note on Dalziel’s model of long run distributive equilibrium: reply’, *Journal of Post Keynesian Economics*, **14**(1), 121–4.

Dasgupta, P. and G.M. Heal (1974), ‘The optimal depletion of exhaustible resources, *Review of Economic Studies*, **41**, 3–28.

Dasgupta, S. and T. Mitra (1988), ‘Intertemporal optimality in a closed linear model of production’, *Journal of Economic Theory*, **45**, 288–315.

Davis, K. and G.J. Moore (1945), ‘Some principles of stratification’, *American Sociological Review*, **10**, 242–9.

Davis, S. and J.C. Haltiwanger (1992), ‘Gross job creation, gross job destruction, and employment reallocation’, *Quarterly Journal of Economics*, **107**(3), 819–64.

de Bruyn, S.M. (2000), *Economic Growth and the Environment: An Empirical Analysis*, Dordrecht: Kluwer.

de Bruyn, S.M. and R.J. Heintz (1999), ‘The environmental Kuznets Curve’, in Van den Bergh J. (ed.), *Handbook of Environmental and Resource Economics*, Aldershot: Edward Elgar, pp. 656–77.

Denicolò, V. and M. Matteuzzi (1990), ‘Public debt and the Pasinetti paradox’, *Cambridge Journal of Economics*, **14**(3), 339–44.

Diamond, D. (1984), ‘Financial intermediation and delegated monitoring’, *Review of Economic Studies*, **51**, 393–414.

Diamond, P. (ed.), (1990), *Growth, Productivity, Unemployment*, Cambridge, Massachusetts: MIT Press.

Dinopoulos E. and P. Thompson (1998), 'Schumpeterian growth without scale effects', *Journal of Economic Growth*, **3**, 313–37.

Dixit, A. and J. Stiglitz (1977), 'Monopolistic competition and optimum product diversity', A*merican Economic Review*, **67**, 297–308.

Dixon, R. and A.P. Thirlwall (1975), 'A model of regional growth-rate differences on Kaldorian lines', *Oxford Economic Papers*, **27**(2), 201–14.

Dolmas, J. (1996), 'Endogenous growth in multisector Ramsey models', *International Economic Review*, **37**, 403–21.

Domar, E. (1946), 'Capital expansion, rate of growth, and employment', *Econometrica*, **14**, 137–47.

Dosi, G. (1982), 'Technological paradigms and technological trajectories', *Research Policy*, **11**, 147–62.

Dosi, G. and R.R. Nelson (1994), 'An introduction to evolutionary theory in economics', *Journal of Evolutionary Economics*, **4**(3), 153–72.

Dow, S. (1985), *Macroeconomic Thought: A Methodological Approach*, Oxford: Basil Blackwell.

Duménil, G. and D. Lévi (1985),'The classicals and the neoclassicals: a rejoinder to Frank Hahn', *Cambridge Journal of Economics,* **9**, 327–45.

Duménil, G. and D. Lévi (1987), 'The dynamics of competition: a restoration of the classical analysis', *Cambridge Journal of Economics,* **11**, 133–64.

Durlauf, S.N. (1996), 'A theory of persistent income inequality', *Journal of Economic Growth*, **1**, 75–93.

Durlauf, S.N. (1999a), 'The case against social capital', *Focus*, **20**, 1–5.

Durlauf, S.N. (1999b), 'Membership theory of inequality', in E. Brezis and P. Temin (eds), *Elites, Minorities and Economic Growth,* Amsterdam: Elsevier Science.

Dutt, A.K. (1984) 'Stagnation, income distribution and monopoly power', *Cambridge Journal of Economics*, **8**(1), 25–40.

Dutt, A.K. (1987), 'Alternative closures again: a comment on 'growth, distribution and inflation', *Cambridge Journal of Economics*, **11**(1), 75–82.

Dutt, A.K. (1990), *Growth, Distribution and Uneven Development*, Cambridge: Cambridge University Press.

Dyson, T. and M. Murphy (1985), 'The onset of fertility transition', *Population and Development Review*, **11**, 399–440.

Eicher, T.S. and S.J. Turnovsky (1999a), 'Non-scale models of economic growth', *Economic Journal*, **109**, 394–415.

Eicher, T.S. and S.J, Turnovsky (1999b), 'Convergence in a two-sector non-scale growth model', *Journal of Economic Growth*, **4**, 413–28.

Eicher, T.S. and S.J. Turnovsky (2001), 'Transition dynamics in a two-sector nonscale rowth model', *Journal of Economic Dynamics and Control*, 25, 85–113.

Eisner, R. (1958), 'On growth models and the neoclassical resurgence', *Economic Journal*, **68**, 707–21.

Eltis, W. (1984), *The Classical Theory of Economic Growth*, London: Macmillan.

Erlich, I. and F.T. Lui (1991), Intergenerational trade, longevity and economic growth, *Journal of Political Economy*, **99**, 1029–59.

Ethier, W. (1982), 'National and international returns to scale in the modern theory of international trade', *American Economic Review*, **72**, 389–405.

Evans, G.W., S. Honkapohja and P.M. Romer (1998), 'Growth cycles', *American Economic Review*, **88**, 495–515.

Fair, R.C. (1999), 'Estimated inflation costs had European unemployment been reduced in the 1980s by macro policies', *Journal of Macroeconomics*, **21**(1), Winter.

Fanti, L. and P. Manfredi (2001), 'Population dynamics, political economy and "old" and "new" growth theory', Working Paper, Dipartimento di Statistica e Matematica Applicata all'Economia, University of Pisa

Fershtman, C., K. Murphy and Y. Weiss (1996), 'Social status, education and growth', *Journal of Political Economy*, **104**, 108–32.

Fisher, S. (1988), 'Recent development in macroeconomics', *Economic Journal*, vol. 98, pp. 294–339.

Flaschel, P. (1988), 'Fiscal policy in an accelerator-augmented classical growth cycle', in P. Flaschel and M. Krüger (eds), *Recent Approaches to Economic Dynamics*, Frankfurt am Main: Peter Lang, pp. 80–99.

Flaschel, P. (1993), *Macrodynamics. Income Distribution*, *Effective Demand and Cyclical Growth*, Frankfurt am Main: Peter Lang.

Flaschel, P., R. Franke and W. Semmler (1997), *Dynamic Macroeconomics. Instability, Fluctuations, and Growth in Monetary Economies*, Cambridge, Massachusetts: MIT Press.

Fleck, F.H. and C.M. Domenghino (1987), 'Cambridge (UK) versus Cambridge (Mass.): a Keynesian solution to the "Pasinetti paradox"', *Journal of Post Keynesian Economics*, **10**(1), 22–36.

Foley, D.K. and T.R. Michl (1999), *Growth and Distribution*, Cambridge, Massachusetts: Harvard University Press.

Follmer, H. (1974), ‘Random economies with many interacting agents’, *Journal of Mathematical Economics*, **1**, 51–62.

Forster, B.A. (1980), ‘Optimal energy use in a polluted environment’, *Journal of Environmental Economics and Management*, **7** (4), 321–33.

Foster, J. (1996), ‘Analytical foundation of evolutionary economics’. *University of Queensland Department of Economics Discussion Papers*, n. 198.

Franke, R.H, G.H. Hofstede and M.H. Bond (1991), ‘Cultural roots of economic performance: a research note’, *Strategic Management Journal*, **12**, 165–73.

Frankel, M. (1962), ‘The production function in allocation and growth: a synthesis’, *American Economic Review*, **52**, 995–1022

Freni, G., F. Gozzi and N. Salvadori (2001), ‘A multisector ‘AK model’ with endogenous growth: existence and characterization of optimal paths and steady state analysis’, *Studi e Ricerche del Dipartimento di Scienze Economiche dell’Università di Pisa*, n. 75, http://www-dse.ec.unipi.it/salvadori/pdf/FGS-web.pdf.

Frisch, R. (1993), ‘Propagation problems and impulse problems in dynamic economics, in *Economic Essays in Honour of Gustav Cassel*, London: Allen and Unwin, pp. 171–205.

Fuà, G. (1993), *Crescita economica: le insidie delle cifre*, Bologna: Il Mulino.

Funk, P. (1996), ‘Relative needs, and the long run economic problem’, unpublished manuscript, Paris: Delta.

Futagami, K. and K. Mino (1993), ‘Threshold externalities and cyclical growth in a stylized model of capital accumulation’, *Economic Letters*, **41**(1), 99–105.

Galbraith, J.K. (1997), ‘Time to ditch the NAIRU’, *Journal of Economic Perspectives*, **11**(1), 93–108.

Gale, D. (1960), *The Theory of Linear Economic Models*, New York: McGraw-Hill.

Gale, D. and M. Hellwig (1985), ‘Incentive-compatible debt contracts: the one-period problem’, *Review of Economic Studie*, **52**, 647–63.

Galeotti, M. and A. Lanza (1999), ‘Desperately seeking (environmental) Kuznets’, *Crenos Working Papers 99/1*, available at http://www.crenos.unica.it/working/99_1.html

Galor, O. and O. Moav (2000), ‘Natural selection and the origin of economic growth’, *Brown University Working Paper*, 2000–18.

Galor, O. and D. Tsiddon (1994), ‘Human capital distribution, technological progress and economic growth’, *CEPR Discussion Paper*, 971.

Galor, O. and D. Tsiddon (1997), ‘Technological progress, mobility and economic growth’, *American Economic Review*, **87**, 363–82.

Galor, O. and D.N. Weil (1996), ‘The gender gap, fertility, and growth’, *American Economic Review*, **86**(3), 374–87.

Galor, O. and D.N. Weil (1998), ‘Population, technology and growth: from the Malthusian regime to the demographic transition’, *NBER Working Paper* n. 6811, November.

Galor, O. and D.N. Weil (1999), ‘From Malthusian stagnation to modern growth’, *American Economic Review*, **89**, 150–4.

Galor, O. and D.N. Weil (2000), ‘Population, technology and growth: from the Malthusian regime to the demographic transition and beyond’, *American Economic Review*, forthcoming.

Gandolfo, G. (1997), *Economic Dynamics. Study Edition*, Berlin: Springer-Verlag.

Garegnani, P. (1990a), ‘Quantity of capital’, in J. Eatwell, M. Milgate and P. Newman (eds), *Capital Theory*, London: Macmillan.

Garegnani, P. (1990b), ‘Sraffa: classical versus marginalist analysis’, in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa: Critical Perspectives on the Revival of Classical Theory*, London: Unwin Hyman, pp. 112–41

Garegnani, P. (1992), ‘Some notes for an analysis of accumulation’, in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady State: A Revival of Growth Theory*, London: Macmillan, pp. 47–72.

Garegnani, P. and A. Palumbo (1998), ‘Accumulation of capital’, in Kurz and Salvadori (1998e), pp. 10–18.

Gellner, E. (1996), *Condition for Liberty*, New York: Penguin Books.

Georgescu-Roegen, N. (1976), *Energy and Economic Myths*, New York and Oxford: Pergamon Press.

Giammanco, M.D. (2002), ‘Competition and technical change in Marx: two perspectives’, *History of Economic Ideas*, forthcoming.

Giampietro, M. and K. Mayumi (2000), ‘Jevons’ paradox, scaling in societal metabolism and the fairy tale of Kuznet curves’, paper presented at *Transitions Towards a Sustainable Europe*, 3rd Biennial Conference of the European Society for Ecological Economics, Vienna, 3–6 May 2000.

Gilbert, G. (1980), ‘Economic growth and the poor in Malthus’ *Essay on Population*’, *History of Political Economy*, **12**, 83–96.

Glaeser, E. and J. Scheinkman (2000), ‘Non-market interactions’, *NBER Working Paper* n. 8053.

Glaeser, E., B. Sacerdote and J. Scheinkman (1996), ‘Crime and social interactions’, *Quarterly Journal of Economics*, **61**, 507–48.

Glaeser, E., D. Laibson and B Sacerdote (2000), ‘The economic approach to social capital’, *NBER Working Paper* n. 7728.

Glombowski, J. and M. Krüger (1988), ‘A Short-period growth cycle model’, *Recherches Economiques de Louvain*, **54**(4), 423–38.

Goldin, C. and L. F. Katz (1999), ‘The returns to skill across the twentieth century United States’, *NBER Working Paper* n. 7126.

Goldsmith, R.Y. (1969), *Financial Structure and Development*, New Haven and London: Yale University Press.

Goodwin, R.M. (1946), ‘Innovations and the irregularity of economic cycles’, *Review of Economic Statistics*, **28**(2), 95–104.

Goodwin, R.M. (1948), ‘Secular and cyclical aspects of the multiplier and the accelerator’, in *Income, Employment and Public Policy. Essays in Honor of Alvin H. Hansen*, New York: W.W. Norton & C., pp. 108–32.

Goodwin, R.M. (1951), ‘The nonlinear accelerator and the persistence of business cycles’, *Econometrica*, **19**(1), 1–17.

Goodwin, R.M. (1953), ‘The problem of trend and cycle’, *Yorkshire Bulletin of Economic and Social Research*, **5**(2), 89–97.

Goodwin, R.M. (1955), ‘A model of cyclical growth’, in E. Lundberg (ed.), *The Business Cycle in the Post-War World. Proceedings of a Conference held in Oxford by the International Economic Association on ‘Business Cycle Problems’, September 1952*, London: Macmillan, pp. 203–21.

Goodwin, R.M. (1967), ‘A growth cycle’, in C.H. Feinstein (ed.), *Socialism, Capitalism and Economic Growth. Essays Presented to Maurice Dobb*, Cambridge: Cambridge University Press, pp. 54–8.

Goodwin, R.M. (1986), ‘The M-K-S-system: the functioning and evolution of capitalism’, in H.-J. Wagener and J.W. Drucker (eds), *The Economic Law of Motion of Modern Society. A Marx-Keynes-Schumpeter Centennial*, Cambridge: Cambridge University Press, pp. 14–21.

Gordon R. (2000), ‘Does the ‘new economy’ measure up to the great inventions of the past?’, *NBER Working Paper* n. W7833.

Gozzi, F. and G. Freni (2001), ‘On a dynamic non-substitution theorem and other issues in Burgstaller’s “Property and Prices”‘, *Metroeconomica*, **52**(2), 181–96.

Gradstein, M. and M. Justman (2000), ‘Human capital, social capital and public schooling’, *European Economic Review*, **44**, 879–90.

Gray, H.P. (1996), ‘Culture and economic performance: policy as an interventing variable’, *Journal of Comparative Economics*, **23**(3), 278–91.

Greenwood, J. and B. Jovanovic (1990), ‘Financial development, growth, and the distribution of income’, *Journal of Political Economy*, **98**, 1076–107.

Greiner, A. and W. Semmler (1996a), ‘Multiple steady states, Indeterminacy, and cycles in a basic model of endogenous growth’, *Journal of Economics*, **63**(1), 79–99.

Greiner, A. and W. Semmler (1996b), ‘Saddle path stability, fluctuations, and indeterminacy in economic growth’, *Studies in Nonlinear Dynamics and Econometrics*, **1**(2), 105–18.

Griliches Z. (1989), ‘Patents: recent trends and puzzles’, *Brooking Papers on Economic Activity, Microeconomics*, 291–330.

Griliches Z. (1990), ‘Patent statistics as economic indicators: a survey’, *Journal of Economic Literature*, **28**, 1612–707.

Grossman, G.M. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge, Massachusetts: MIT Press.

Grossman, G.M. and E. Helpman (1994), ‘Endogenous innovation in the theory of growth’, *Journal of Economic Perspectives*, **8**(1), 23–44.

Grossman, H.I. and M. Kim (1995), ‘Swords or plowshares? a theory of the security of claims to property’, *Journal of Political Economy*, **103**, 1275–88.

Gurley, J.G. and E.S. Shaw (1955), ‘Financial aspects of economic development’, *American Economic Review*, **45**, 515–38.

Gurley, J.G. and E.S. Shaw (1960), *Money in a Theory of Finance*, Washington, DC: Brookings Institute.

Gurley, J.G. and E.S. Shaw (1967), ‘Financial structure and economic development’ *Economic Development and Cultural Change*, **15**, 257–68.

Hagemann, H. (1998), ‘General glut controversy?, in Kurz and Salvadori (1998e), pp. 320–4.

Hahn, F.H. and R. Matthews (1964), ‘The theory of economic growth: a survey’, *Economic Journal*, **74**, 779–901, reprinted in *Surveys of Economic Theory, Growth and Development*, London: Macmillan, 1965.

Hall, R.E. (1993), ‘Macro theory and the recession of 1990–91’, *American Economic Review*, **83**(2), 275–9.

Hall, R.E. (1999), ‘Reorganizations’, *NBER Working Paper* n. 7181.

Hamberg, D. (1971), *Models of Economic Growth*, New York: Harper and Row.

Hamouda, O. and G. Harcourt (1989), ‘Post-Keynesianism: from criticism to coherence?’, in J. Pheby (ed.), *New Directions in Post-Keynesian Economics*, Aldershot: Edward Elgar, pp. 1–34.

Hands, D.W. (1988), ‘Ad hocness in economics and the Popperian tradition’, in N. De Marchi (ed.), *The Popperian Legacy in Economics*, Cambridge: Cambridge University Press, pp. 121–37.

Hands, D.W. (1990), ‘Second thoughts on “second thoughts”: reconsidering the Lakatosian progress of *The General Theory*’, *Review of Political Economy*, **2**, 69–81.

Hansen, G.D. and E.C. Prescott (1999), *From Malthus to Solow*, Federal Reserve Bank of Minneapolis, Research Department Staff Report 257.

Harcourt, G.C. (1972), *Some Capital Controversies in the Theory of Capital.* Cambridge: Cambridge University Press.

Harris, D.J. (1980), ‘A postmortem on the neo-classical parable’, in E.J. Nell (ed.) *Growth, Profits, and Property*, Cambridge: Cambridge University Press, pp. 43–63.

Harrod, R.F. (1933), *International Economics*, Cambridge: Cambridge University Press.

Harrod, R.F. (1936), *The Trade Cycle: An Essay*, Oxford: Clarendon Press.

Harrod, R.F. (1939), ‘An essay in dynamic theory’, *Economic Journal*, **49**(193), 14–33. Reprinted in R.F. Harrod (1972) *Economic Essays*, 2nd edition, London: Macmillan.

Harrod, R.F. (1948), *Towards a Dynamic Economics*, London: Macmillan.

Harrod, R.F. (1951), ‘Notes on trade cycle theory’, *Economic Journal*, **61**(242), 261–75.

Harrod, R.F. (1951), *The Life of John Maynard Keynes*, London: Macmillan.

Harrod, R.F. (1960), ‘A second essay in dynamic theory’, *Economic Journal*, **70**, 277–93.

Harrod, R.F. (1964), ‘Are monetary and fiscal policies enough?’, *Economic Journal*, **74**, 903–15.

Harrod, R.F. (1967), ‘Increasing returns’, in R. Kuenne (ed.), *Monopolistic Competition: Essays in Honour of Edward Chamberlain*, London: John Wiley and Sons,. Reprinted in R.F. Harrod, *Economic Essays*, London: Macmillan, 1972, pp. 302–17.

Harrod, R.F. (1969), *Money*, London: Macmillan.

Harrod, R.F. (1973), *Economic Dynamics*, London: Macmillan.

Hassler, J. and J. Rodriquez Mora (2000), ‘Intelligence, social mobility, and growth’, *American Economic Review*, **90**(4), 888–908.

Heal, G. (1999), ‘Introduction’, in G. Heal (ed.), *The Economics of Increasing Returns*, Cheltenham: Edward Elgar, pp. xiii–xxvii.

Helpman E. and M. Trajtenberg (1994), ‘A time to sow and a time to reap: growth based on general purpose technologies’, *CEPR Working Paper*, n. 1080.

Hicks, J.R. (1950), *A Contribution to the Theory of the Trade Cycle*, Oxford: Clarendon Press.

Hicks, J. (1969), *A Theory of Economic History*, Oxford: Clarendon Press.

Hirsh, F. (1976), *Social Limits to Growth*, Cambridge, Massachusetts: Harvard University Press.

Hollander, S. (1973), *The Economics of Adam Smith*, London: Heinemann Educational Books.

Houthakker, H. and S. Magee (1969), ‘Income and price elasticities in world trade’, *Review of Economics and Statistics*, **51**(2), 111–25.

Howitt P. (1999), ‘Steady endogenous growth with population and R&D growing’, *Journal of Political Economy*, **107**, 715–30.

Iwai, K. (2000), ‘A contribution to the evolutionary theory of innovation, imitation and growth’, *Journal of Economic Behavior and Organization*, **43**, 167–98.

Iyigun, M.F., H.N. Mocan and A.L. Owen (2001), ‘Ideology, human capital and growth’, unpublished manuscript, University of Colorado.

Jones, C.I. (1995), ‘R&D based models of economic growth’, *Journal of Political Economy*, **103**, 759–84.

Jones, C.I. (1998a), *Introduction to Economic Growth*, New York and London; W.W. Norton.

Jones, C.I. (1998b), *Population and Ideas: A Theory of Endogenous Growth*, Department of Economics, Stanford University, *mimeo.*

Jones, C.I. (1999a), ‘Growth: with or without scale effects?’, *American Economic Review*, **89**(2), 139–44.

Jones, C.I. (1999b), ‘Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run’, *NBER Working Paper* n. 7375.

Jones, C.I. (2002), “Sources of economic growth in a world of ideas”, *American Economic Review*, 92, 220–239.

Jones, L. and R. Manuelli (1990), ‘A convex model of equilibrium growth: theory and policy implications’, *Journal of Political Economy*, **98**, 1008–38.

Jones, L.E. and R. Manuelli (1997), ‘The sources of growth’, *Journal of Economic Dynamics and Control*, **21**, 75–114.

Jorgenson, D.W. (1967), ‘The theory of investment behavior’, in R. Ferber (ed.), *Determinants of investment behaviour*, New York: NBER, Columbia University Press.

Jovanovich B. and R. Rob (1990), ‘Long waves and short waves: growth through intensive and extensive search’, *Econometrica*, **58**, 1391–409.

Junankar, P.N. (1972), *Investment: Theories and evidence*, London: Macmillan.

Kaganovich, M. (1998), ‘Sustained endogenous growth with decreasing returns and heterogeneous capital’, *Journal of Economic Dynamics and Control*, **22**, 1575–603.

Kaldor, N. (1954a), ‘The relation of economic growth and cyclical fluctuation’, *Economic Journal*, **64**(253), 53–71, reprinted in Kaldor (1980), pp..

Kaldor, N. (1954b), ‘Characteristics of economic development’, reprinted in Kaldor (1980),pp. 233–42.

Kaldor, N. (1955–56), ‘Alternative theories of distribution’, *Review of Economic Studies*, **23**(2), 83–100.

Kaldor, N. (1957a), ‘A model of economic growth’, *Economic Journal*, **67**, 591–624.

Kaldor, N. (1957b), ‘Capitalist evolution in the light of Keynesian economics’, *Sankhya* (May 1957)???, reprinted in Kaldor (1980), pp. 243–58.

Kaldor, N. (1958), ‘Monetary policy, economic stability and growth: a memorandum submitted to the Radcliffe Committee on the working of the monetary system’, 23, *Principal Memoranda of Evidence*, Cmnd 827, H.M.S.O, London, 1960, 146–53, reprinted in N. Kaldor, *Essays on the Economic Policy I*, London: Duckworth.

Kaldor, N. (1961), ‘Capital accumulation and economic growth’, in F.A. Lutz and D.C. Hague (eds), *The Theory of Capital*, London: Macmillan, pp. 177–222.

Kaldor, N. (1966), *Causes of the Slow Rate of Growth of the United Kingdom: An Inaugural Lecture*, Cambridge: Cambridge University Press, reprinted in Kaldor (1978a).

Kaldor, N. (1967), *Strategic Factors in Economic Development*, New York: Ithaca, Cornell University.

Kaldor, N. (1970), ‘The case for regional policies’, *Scottish Journal of Political Economy*, **17**, 337–48, reprinted in Kaldor (1978a).

Kaldor, N. (1971), ‘Conflicts in national economic objectives’, *Economic Journal*, **81**, 1–16, reprinted in Kaldor (1978a).

Kaldor, N. (1972), ‘The irrelevance of equilibrium economics, *Economic Journal*, **82**, 1237–55.

Kaldor, N. (1975), ‘What is wrong with economic theory’, *Quarterly Journal of Economics*, **89**, 347–57.

Kaldor, N. (1978a), *Further Essays on Economic Theory*, London: Duckworth.

Kaldor, N. (1978b), *Further Essays on Applied Economics*, London: Duckworth.

Kaldor, N. (1978c), ‘The effects of devaluation on trade in manufactures’, in Kaldor (1978b).

Kaldor, N. (1980), *Essays on Economic Stability and Growth*, London: Duckworth.

Kaldor, N. (1981), ‘The role of increasing returns, technical progress and cumulative causation in the theory of international trade and economic growth’, *Economie Appliquée*, **34**, 593–617.

Kalecki, M. (1937), ‘The principle of increasing risk’, *Economica*,**4**(16), 440–447.

Kalecki, M. (1968), ‘Trend and business cycles reconsidered’, *Economic Journal*, **78**(310), 263–76.

Kalecki, M. (1971), *Selected Essays in the Dynamics of the Capitalist Economy*, Cambridge: Cambridge University Press.

Keeton, W. (1979), *Equilibrium Credit Rationing*, New York: Garland Publishing Company.

Keynes, J.M. (1973–89), *The Collected Writings of J.M. Keynes*, 30 vols, edited by D. Moggridge, London: Macmillan. Referred to in the text as *CW*, **volume** number, page number.

King, R.G. and R. Levine (1993a), ‘Finance and growth: Schumpeter might be right’, *Quarterly Journal of Economics*, **108**, 717–37.

King, R.G. and R. Levine (1993b), ‘Finance entrepreneurship, and growth: theory and evidence’, *Journal of Monetary Economics*, **32**, 513–42.

King, R.G. and S. Rebelo (1990), ‘Public policy and economic growth: developing neoclassical implications’, *Journal of Political Economy*, **98**, 126–50.

King, R.G., C.I. Plosser and S.T. Rebelo (1988), ‘Production, growth and business cycles. I. The basic neoclassical model’, *Journal of Monetary Economics*, **21**(2–3), 195–232.

Kirman, A. (1999), ‘Interaction and markets’, in M. Gallegati and A. Kirman (eds) *Beyond the Representative Agent*, Chelthenam: Edward Elgar Publishing, pp. 1–44.

Knack, S. and P. Keefer (1997), ‘Does social capital have an economic payoff? A cross-country investigation’, *Quarterly Journal of Economics*, **112**, 1252–88.

Knodel, J.A. (1977), ‘Age patterns of fertility and the fertility transition: evidence from Europe and Asia’, *Population Studies*, **31**(2), 219–50.

Kögel, T. and A. Prskawetz (2001), ‘Agricultural Productivity Growth and Escape from the Malthusian Trap’, *Journal of Economic Growth*, **6**(4), 337–57.

Kongsamut P., S. Rebelo, and D. Xie (2001), ‘Beyond balanced growth’, *Review of Economic Studies*, **68(**4), 869–82.

Koopmans, T.C. (1965), ‘On the concept of optimal economic growth’, in *The Econometric Approach to Development Planning*, Amsterdam: North-Holland, pp. 225–87.

Kortum S.S. (1997), 'Research, patenting and technological change', *Econometrica*, **65**, 1389–419.

Kregel, J.A. (1980), 'Economic dynamics and the theory of steady growth: an historical essay on Harrod's "knife-edge"', *History of Political Economy*, **12**(1), 97–123.

Kregel, J.A. (1985), 'Harrod and Keynes: increasing returns, the theory of employment, and dynamic economics', in G.C. Harcourt (ed.), *Keynes and his Contemporaries*, London: Macmillan, pp. 66–88.

Kremer, M. (1993), 'Population growth and technological change: one million B.C. to 1990', *Quarterly Journal of Economics*, **108**, 681–716.

Krueger, A.O. (1974), 'The political economy of the rent-seeking society', *American Economic Review*, **64**, 291–303.

Kurz, H.D. (1986), 'Normal positions and capital utilization', *Political Economy*, **2**(1), 37–54.

Kurz, H.D. (1991), 'Technical change, growth and distribution: a steady state approach to unsteady growth on Kaldorian lines', in Nell and Semmler (1991), pp. 421–48.

Kurz, H.D. (1992), 'Accumulation, effective demand and income distribution', in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady State: A Revival of Growth Theory*, London: Macmillan, pp. 73–95.

Kurz, H.D. and N. Salvadori (1993), 'Von Neumann's growth model and the "classical" tradition', *The European Journal of the History of Economic Thought*, **1**, 129–60.

Kurz, H.D. and N. Salvadori (1994), 'The non-substitution theorem: making good a lacuna', *Journal of Economics*, **59**, 97–103.

Kurz, H.D. and N. Salvadori (1995), *Theory of Production: A Long-Period Analysis*, Cambridge: Cambridge University Press.

Kurz, H.D. and N. Salvadori (1998a), '"Endogenous" growth models and the "classical" tradition', in Kurz and Salvadori (1998d), pp. 66–89.

Kurz, H.D. and N. Salvadori (1998b), 'Postscript (to 'The non-substitution theorem: making good a lacuna') in Kurz and Salvadori (1998d), pp. 95–9.

Kurz, H.D. and N. Salvadori (1998c), *Understanding "Classical" Economics. Studies in the Long-Period Theory*, London and New York: Routledge.

Kurz, H.D. and N. Salvadori (eds), (1998d), *The Elgar Companion to Classical Economics*, vol. I, Chelthenam: Edward Elgar.

Kurz, H.D. and N. Salvadori (1999), 'Theories of "endogenous" growth in historical perspective', in M.R. Sertel (ed.) *Contemporary Economic Issues. Proceedings of the Eleventh World Congress of the International Economic Association, Tunis*, vol. 4: 'Economic Behaviour and Design', London: Macmillan, pp. 225–61.

Kydland, F.E. and E.C. Prescott (1982), 'Time to build and aggregate fluctuations', *Econometrica*, **50**(6), 1345–70.

Ladron-de-Guevara, A., S. Ortigueira and M.S. Santos (1999), 'A two-sector model of endogenous growth with leisure', *Review of Economic Studies*, **66**, 609–31.

Laitner, J. (2000), 'Structural change and economic growth', *Review of Economic Studies*, **67**, 545–61.

Lakatos, I. (1978), *The Methodology of Scientific Research Programmes*, Cambridge: Cambridge University Press.

Landes, D.S. (1969). *The Unbound Prometheus*. Cambridge: Cambridge University Press.

Lanjow J.O. and M. Schankerman (1999), 'The quality of ideas: measuring innovation with multiple indicators', *NBER Working Paper* n. 7345.

Latsis, S.J. (1976), *Method and Appraisal in Economics*, Cambridge: Cambridge University Press.

Lavezzi, A. (2002), ' Smith, Marshall and Young on Division of Labor and Economic Growth', *European Journal of the History of Economic Thought*, forthcoming.

Lavoie, M. (1992), *Foundations of Post-Keynesian Economic Analysis*, Aldershot: Edward Elgar.

Lavoie, M. (1995), 'The Kaleckian model of growth and distribution and its neo-Ricardian and Neo-Marxian critiques', *Cambridge Journal of Economics*, **19**(6), 789–818.

Lazear, E.P. (1999), 'Culture and language', *Journal of Political Economy*, **107**, S95–1126.

Lee, R.D. (1986), 'Malthus and Boserup: a dynamic synthesisp', in D. Coleman and R. Schofield (eds), *The State of Population Theory: Forward from Malthus*, Oxford: Basil Blackwell, pp. 96–130.

Lee, R.D. (1997), 'Population dynamics: equilibrium, disequilibrium, and consequences of fluctuations', in M.R. Rosenzweig and O. Stark (eds), *Handbook of Population and Family Economics*, Amsterdam: North-Holland, pp. 1063–116.

Lerner, A.P. (1944), *The Economics of Control*, London: Macmillan.

Levhari, D. (1966), 'Extensions of Arrow's learning by doing', *Review of Economic Studies*, **33**(1), 117–31.

Levine R. (1991), 'Stock markets, growth and tax policy', *The Journal of Finance*, **46**, 1445–65.

Levine, R. (1997), 'Financial development and economic growth: views and agenda', *Journal of Economic Literature*, **35**, 688–726.

Levine, R. and S. Zervos (1996), 'Stock market development and long-run growth', *The World Bank Economic Review*, **10**(2), 223–39.

Levine, R. and S. Zervos (1998), 'Stock Markets, Banks and Economic Growth', *American Economic Review*, **88**, 537–57.

Lindbeck, A. and D.J. Snower (1999), 'Price dynamics and production lags', *American Economic Review*, **89(**2), May.

Lippi, M. (1993), 'Sviluppi recenti del modello lineare-stocastico in macroeconomia, con particolare attenzione alle teorie del ciclo reale', in A. Medio (ed.), *Analisi dinamica in economia: modelli non-lineari e lineari stocastici*, Bologna: Il Mulino, pp. 35–52.

Livi-Bacci, M. (1997). *A Concise History of World Population*, second edition, Oxford: Blackwell.

Livi-Bacci, M. (1998), *La popolazione nella storia d'Europa*, Bari: Laterza.

Long, J. and C. Plosser (1983), 'Real business cycles', *Journal of Political Economy*, **91**(1), 39–69.

Lowe, A. (1954), 'The classical theory of economic growth', *Social Research*, **21**, 127–58, reprinted in A. Lowe, *Essays in Political Economics: Public Control in a Democratic Society*, edited by A. Oakley, Brighton: Wheatsheaf Books, 1987.

Lowe, A. (1975), 'Adam Smith's Theory of Equilibrium Growth', in Skinner and Wilson (1975), pp. 415–25.

Lucas, R.E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, **22**(1), 3–42.

Lucas, R.E. (1993), 'Making a miracle', *Econometrica*, **61**, 251–72.

Lucas, R.E. (1997), 'On the mechanics of economic development', in D.P. Jacobs, E. Kalai, M.I. Kamien (eds), *Frontiers on Research in Economic Theory*, Cambridge: Cambridge University Press, pp. 61–71.

Lucas, R.E. (1998), *The Industrial Revolution: Past and Future*, Unpublished Manuscript, University of Chicago.

Maccabelli, T. (1997). *Il progresso della ricchezza. Economia, politica e religione in T.R. Malthus*, Milano: Giuffrè Editore.

Malinvaud, E. (1995), 'About investment in macroeconomics', in J.-P. Fitoussi (ed.), *Economics in a Changing World, vol. 5: Economic Growth and Capital and Labour Markets*, New York: St. Martin's Press, pp. 123–41.

Malthus, T.R. (1986). *The Works of Thomas Robert Malthus*, edited by E.A. Wrigley and D. Souden, London: William Pickering.

Manski, C.F. (2000), 'Economic analysis of social interactions'', *Journal of Economic Perspectives*, **14**, 115–36.

Marglin, S.A. (1984a), *Growth, Distribution and Prices*, Cambridge, Massachusetts: Harvard University Press.

Marglin, S.A. (1984b), 'Growth, distribution, and inflation: a centennial synthesis', *Cambridge Journal of Economics*, **8**(2), 115–44.

Marshall, A. (1890), *Principles of Economics*, 8th edn 1920. Reprint, reset (1977), London: Macmillan.

Marshall, M.G. (2000), 'Luxury, economic development, and work motivation: David Hume, Adam Smith, and J.R. McCulloch', *History of Political Economy*, **32**, 631–48.

Martinez, A. (1987), *Ecological Economics: Energy, Environment and Society*, Oxford: Blackwell.

Marx, K. (1954), *Capital*, vol. I, Moscow: Progress Publishers. English translation of *Das Kapital*, vol. I (1867), Hamburg: Meissner.

Marx, K. (1956), *Capital*, vol. II, Moscow: Progress Publishers. English translation of *Das Kapital*, vol. II, edited by F. Engels (1885), Hamburg: Meissner.

Mattana, P. and B. Venturi (1999), 'Existence and stability of periodic solutions in the dynamics of endogenous growth', *Rivista internazionale di Scienze Economiche e Sociali*, **46**(2), 259–84.

McCombie, J.S.L. (1998), 'Harrod, economic growth and international trade', in G. Rampa, L. Stella, A.P. Thirlwall (eds), *Economic Dynamics, Trade and Growth. Essays on Harrodian Themes*, London: Macmillan Press, pp. 212–44.

McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance-of-Payments Constraint*, New York: St.Martin's Press.

McCombie, J.S.L. and A.P. Thirlwall (1999), 'Growth in an international context: a post Keynesian view', in J. Deprez and J.T. Harvey (eds), *Foundations of International Economics. Post Keynesian Perspectives*, London: Routledge, pp. 35–90.

McFadden, D. (1967), 'The evaluation of development programmes', *Review of Economic Studies*, **34**, 25–50.

McGrattan, E.R. (1990), 'Solving the stochastic growth model by linear-quadratic approximation', *Journal of Business & Economic Statistics*, **8**(1), 41–4.

McKenzie, L.W. (1998), 'Turnpikes', *American Economic Review*, **88**, 1–14.

McKinnon, R.I. (1973), *Money and Capital in Economic Development*, Washington, D.C.: Brookings Institution.

McNulty, P.J. (1967), ‘A note on the history of competition’, *Journal of Political Economy,* **75**, 396–9.

McNulty, P.J. (1968), ‘Economic theory and the meaning of competition’, *Quarterly Journal of Economics,* **82**, 639–56.

Meade, J.E. (1961), *A Neoclassical Theory of Economic Growth*, London: Allen and Unwin.

Meadows, D.H., D.L. Meadows, J. Randers and W. Behrens III (1972), *The Limits to Growth*, New York: Universe Book.

Merton R. (1965), *On the Shoulders of Giants: A Shandean Postscript*, New York: Free Press.

Mirrlees, J. and Stern, N. (1974) (eds), *Models of economic growth,* London: Macmillan.

Mohtadi, H. and T. Roe (1998), ‘Growth, lobbying and public goods’, *European Journal of Political Economy*, **14**, 453–73.

Momota, A. and K. Futagami (2000), ‘Demographic transition pattern in a small country’, *Economic Letters*, **67**, 231–7.

Moreno Brid, J.C. (1998–99), ‘On capital flows and the balance-of-payments constrained growth model’, *Journal of Post Keynesian Economics,* **21**(2), 283–98.

Moss, S.J. (1978), ‘The post Keynesian theory of income distribution in the corporate economy’, *Australian Economic Papers*, **38**(4), 303–22.

Mulligan, C.B. and X. Sala-i-Martin (1993), ‘Transitional dynamics in two-sector models of endogenous growth’, *Quarterly Journal of Economics,* **108**(3), 739–73.

Murphy K., A. Shleifer, and R. Vishny (1989a), ‘Income distribution, market size, and industrialization’, *Quartely Journal of Economics*, **104**, 537–64.

Murphy K., A. Shleifer and R. Vishny (1989b), ‘Industrialization and the big push’, *Journal of Political Economy*, **97**, 1003–26.

Murphy, K.D., A. Shleifer and R. Vishny (1991), ‘The allocation of talent: implication for growth’, *Quarterly Journal of Economics,* **106**, 503–30.

Murphy, K.D., A. Shleifer and R. Vishny (1993), ‘Why is rent-seeking so costly to growth’, *American Economic Review, Papers and Proceedings*, **83**, 409–14

Negishi, T. (1993), ‘A Smithian growth model and Malthus’s optimal propensity to save’, *The European Journal of the History of Economic Thought,* **1**, pp. 115–27.

Nell, E.J. (1985), ‘Jean Baptiste Marglin: a comment on “Growth, distribution and inflation”‘, *Cambridge Journal of Economics*, **9**(2), 173–8.

Nell, E.J. (1988), ‘Does the rate of interest determine the rate of profit?’ *Political Economy*, **4**(2), 263–7.

Nell, E.J. and W. Semmler (1991) (eds), *Nicholas Kaldor and the Mainstream Economics*, London: Macmillan.

Nelson, C. and C. Plosser (1982), ‘Trends and random walks in macroeconomic time series: some evidence and implications’, *Journal of Monetary Economics,* **10**(2), 139–62.

Nelson, R.R. (1956), ‘A theory of the low level equilibrium trap in under-developed economies’, *American Economic Review*, **46**, 894–908.

Nelson, R.R. (1973), ‘Recent exercises in growth accounting: new understanding or dead end?’, *American Economic Review*, **63**(3), 462–8.

Nelson, R.R. (1994), ‘What has been the matter with neoclassical growth theory’, in G. Silverberg and L. Soete (eds), *The Economics of Growth and Technical Change*, Aldershot: Edward Elgar, pp. 290–334.

Nelson, R.R. (1996), *The Source of Economic Growth*, Cambridge, Massachusetts: Harvard University Press.

Nelson, R.R. (1998), ‘The agenda for growth theory: a different point of view’, *Cambridge Journal of Economics*, **22**, 497–520.

Nelson, R.R. and S.G. Winter (1974), ‘Neoclassical vs. evolutionary theories of economic growth: critiques and prospectus’, *Economic Journal*, **84**, 886–905.

Nelson, R.R. and S.G. Winter (1977), ‘In search of a useful theory of innovation’, *Research Policy*, **6**, 37–76.

Nelson, R.R. and S.G. Winter (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Massachusetts: Harvard University Press.

Nelson, R.R., S.G. Winter and L.H. Schuette (1976), ‘Technological change in an evolutionary model’, *Quarterly Journal of Economics*, **90**, 90–118.

Nerlove, M. and L.K. Raut (1997), ‘Growth models with endogenous population: a general framework’, in M.R. Rosenzweig and O. Stark (eds), *Handbook of Population and Family Fconomics*, Amsterdam: North-Holland, pp. 1117–74.

Niehans, J. (1963), ‘Economic growth with two endogenous factors’, *Quarterly Journal of Economics*, **77**, 349–71.

Nordhaus W. (1997), ‘Do real-output and real-wage measures capture reality? the history of light suggests not’, in T. Bresnahan and R.J. Gordon (eds), *The Economics of New Goods*, Chicago: University of Chicago Press, pp. 29–66.

North, D. (1981), *Structure and Change in Economic History*, New York: W.W. Norton and Company, Inc.

Ossella, I. (1999), ‘A turnpike property of optimal programs for a class of simple linear models of production’, *Economic Theory*, **14**, 597–607.

Palivos, T. (1995), ‘Endogenous fertility, multiple growth paths, and economic convergence’, *Journal of Economic Dynamics and Control*, **19**, 1489–1510.

Panico, C. (1993), ‘Two alternative approaches to financial model building’, *Metroeconomica*, **44**(2), 93–133.

Panico, C. (1997), ‘Government deficits in the post keynesian theories of growth and distribution’, *Contributions to Political Economy*, **16**, 61–86.

Panico, C. (1999), ‘The government sector in the post keynesian theory of growth and personal distribution’, in G. Mongiovi and F. Petri (eds), *Value, Distribution and Capital, Essays in Honour of P. Garegnani*, London: Routledge, pp. 339–53.

Panico, C. (2001), ‘Monetary analysis in Sraffa’s writings’, in T. Cozzi and R. Marchionatti (eds), *Piero Sraffa’s Political Economy: A Centenary Estimate*, London: Routledge, pp. 285–310.

Panico, C. and N. Salvadori (eds), (1993), *Post Keynesian Theory of Growth and Distribution*, Aldershot: Edward Elgar.

Park, M.-S. (2000), ‘Autonomous demand and the warranted rate of growth’, *Contributions to Political Economy*, **19**, 1–18.

Pasinetti, L.L. (1960a), ‘A mathematical formulation of the Ricardian system’, *Review of Economic Studies*, **27**(2), 78–98.

Pasinetti, L.L. (1960b), ‘Cyclical fluctuations and economic growth’, *Oxford Economic Papers*, **12**(2) 18–50.

Pasinetti, L.L. (1962), ‘Rate of profit and income distribution in relation to the rate of economic growth’, *Review of Economic Studies*, **29**(4), 103–20.

Pasinetti, L.L. (1981), *Structural Change and Economic Growth: A Theoretical Essay on the Dynamics of the Wealth of Nations*, Cambridge: Cambridge University Press.

Pasinetti, L.L. (1988), ‘Sraffa on income distribution’, *Cambridge Journal of Economics*, **12**(1), 135–8.

Pasinetti, L.L. (1989a), ‘Government deficit spending is not incompatible with the Cambridge theorem of the rate of profit: a reply to Fleck and Domenghino’, *Journal of Post Keynesian Economics*, **11**(4), 640–7.

Pasinetti, L.L. (1989b), ‘Ricardian debt-taxation equivalence in the Kaldor theory of profits and income distribution’, *Cambridge Journal of Economics*, **13**(1), 25–36.

Pasinetti, L.L. (1993). *Structural Economic Dynamics. A Theory of Economic Consequences of Human Learning*. Cambridge: Cambridge University Press.

Patinkin, D. (1987), entry ‘Real Balances’, in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave Dictionary of Economics*, London: Macmillan.

Pecorino, P. (1992), ‘Rent seeking and growth: the case of growth through human capital’, *Canadian Journal of Economics*, **25**, 944–56.

Peretto P.F. (1998), ‘Technological change and population growth’, *Journal of Economic Growth*, **3**, 283–311.

Perez, C. (1985), ‘Microelectronics, long waves and world structural change: new perspective for developing countries’, *World Development*, **13**, 441–63.

Petith, H. (2001), ‘Georgescu-Roegen versus Solow/Stiglitz and the convergence to the Cobb-Douglas’, Working Paper 481.01 of the Departament d’Economia i d’Història Economíca, Universitat Autonoma de Barcellona.

Petri, F. (1997), ‘On the theory of aggregate investment as a function of the rate of interest’, *Quaderni del Dipartimento di Economia Politica*, Università di Siena, no. 215 (http://www.econ-pol.unisi.it/quaderni/215.pdf).

Phelps Brown, H. (1980), ‘Sir Roy Harrod: a biographical memoir’, *Economic Journal*, **90**, 1–33.

Phelps, E. (1962), ‘The new view of investment: a neo-classical analysis’, *Quarterly Journal of Economics*, **76**, 548–67.

Phillips, A.W. (1954), ‘Stabilisation policy in a closed economy’, *Economic Journal*, **64**(254), 290–323.

Pivetti, M. (1988), ‘On the monetary explanation of distribution: a rejoinder to Nell and Wray’, *Political Economy*, **4**(2), 275–83.

Pivetti, M. (1991), *An Essay on Money and Distribution*, London: Macmillan.

Pomini, Mario (2000), ‘Some notes on the early neo-classical growth models’, *History of Economics Ideas*, **8**, 123–41.

Porter M.E. and S. Stern (2000), ‘Measuring the ‘ideas’ production function: evidence from international patent output’, *NBER Working Paper* n. 7891.

Postlewaite, A. (1998), ‘Social status, norms and economic performance’, *European Economic Review*, **42**, 779–800.

Prskawetz, A., G. Steinmann G. and G. Feichtinger (2000), ‘Human capital, technological progress and the demographic transition’, *Mathematical Population Studies*, **7**(4), 343–63.

Putnam, R. (1993), ‘The prosperous community: social capital and public life’, *The American Prospect*, **13**, 35–42.

Quah, D. (1999), ‘The weightless economy in economic development’, *Centre for Economic Performance Discussion Paper* n. 417, London School of Economics

Rama, M. (1993), ‘Rent seeking and economic growth. a theoretical model and some empirical evidence’, *Journal of Development Economics*, **42**, 35–50.

Ramanathan, R. (1982), *Introduction to a Theory of Economic Growth*, Berlin: Springer-Verlag.

Rampa, G. and L. Rampa (1988), 'Fluctuating growth with induced innovations', *Metroeconomica*, **39**(1), 31–42.

Ramsey, F.P. (1928), 'A mathematical theory of saving', *Economic Journal*, **38**, 543–59.

Razin, A. and U. Ben-Zion (1975), 'An intergenerational model of population growth', *American Economic Review*, **65**(5), 923–33.

Rebelo, S. (1991), 'Long run policy analysis and long run growth', *Journal of Political Economy*, **99**, 500–21.

Remenyi J., (1979), 'Core demi-core interaction: toward a general theory of disciplinary and subdisciplinary growth', *History of Political Economy*, **11**(1), 31–61.

Retherford, R.D. and J.A. Palmore (1983), 'Diffusion processes affecting fertility regulation', in R.A. Bulatao and R.D. Lee (eds), *Determinants of Fertility in Developing Countries*, vol. 2, New York: Academic Press, pp. 295–339.

Ricardo, D. (1951–73). *The Works and Correspondence of David Ricardo*, edited by Piero Sraffa with the collaboration of Maurice H. Dobb, 11 Vols, Cambridge: Cambridge University Press. Referred to in the text as *Works*, **volume** number, page number.

Richardson, G.B. (1975), 'Adam Smith on competition and increasing returns', in Skinner and Wilson (1975), pp. 350–360.

Ricoy, C. (1987), 'Cumulative causation', in J.L. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave. A Dictionary of Economics*, vol. I, London: Macmillan, pp. 730–36.

Ricoy, C. (1998), 'Cumulative causation', in Kurz and Salvadori (1998e), pp. 202–8.

Robinson, J. (1956), *The Accumulation of Capital*, London: Macmillan.

Robinson, J. (1962), *Essays in the Theory of Economic Growth*, London: Macmillan.

Rochon, L.P. (1999), *Credit, Money and Production*, Chelthenam: Edward Elgar.

Romer, D. (2000), 'Keynesian macroeconomics without the LM Curve', *Journal of Economic Perspectives*, **14**(2), 149–69.

Romer, P.M. (1986b), 'Cake eating, chattering and jumps: existence results for variational problems', *Econometrica*, **54**(4), 897–908.

Romer, P.M. (1986c), 'Increasing returns and long-run growth', *Journal of Political Economy*, **94**(5), 1002–37.

Romer, P.M. (1987), 'Growth based on increasing returns due to specialization', *American Economic Review*, **77**, 56–62.

Romer, P.M. (1989), 'Capital accumulation in the theory of long-run growth', in R. Barro (ed.), *Modern Business Cycle Theory*, Cambridge, Massachusetts: Harvard University Press, pp. 51–127.

Romer, P.M. (1990a), 'Are non convexities important to understanding growth?', *NBER Working Paper* n. 3089.

Romer, P.M. (1990b), 'Endogenous technological change', *Journal of Political Economy*, **98**, 71–102.

Romer, P.M. (1991), 'Increasing returns and new developments in the theory of growth', in W.A. Barnett, B. Cornet and C. d'Aspremont (eds), *Equilibrium Theory and Applications*, Cambridge: Cambridge University Press, pp. 83–110.

Romer, P.M. (1994), 'The origins of endogenous growth', *Journal of Economic Perspectives*, **8**(1), 3–22.

Rosenberg, N. (1968), 'Adam Smith, consumer tastes, and economic growth', *Journal of Political Economy*, **76**, 361–74.

Rosenberg, N. (1975), 'Adam Smith on profits – paradox lost and regained', in Skinner and Wilson (1975), pp. 377–89.

Rosenberg, N. (1982), *Inside the Black Box: Technology and Economics*, Cambridge: Cambridge University Press.

Rosenberg, N. and L.E. Birdzell Jr. (1986), *How the West Grew Rich*, New York: Basic Books.

Rosero-Bixby, L. and J.H. Casterline (1993), 'Modelling diffusion effects in fertility transition', *Population Studies*, **47**, 147–67.

Rowley, C.K., R.D. Tollison and G. Tullock (1988), *The Political Economy of Rent-Seeking*, Boston: Kluwer.

Rowley, R. (1995), 'History, structure and the wandering natural rate of unemployment', *Economie Appliquée*, **48**(1), 133–55.

Rowthorn, R.E. (1975), 'A reply to Lord Kaldor's comment', *Economic Journal*, **85**, 897–901.

Rowthorn, R.E. (1981), 'Demand, real wages and economic growth', *Thames Papers in Political Economy*, Autumn, 1–39.

Ryder, H. (1967), 'Optimal accumulation and trade in a open economy of moderate size', in K. Shell (ed.), *Essays on the Theory of Optimal Economic Growth*, Cambridge, Massachusetts: MIT Press, pp. 87–117.

Sacerdote, B. and E.L. Glaeser (2001), 'Education and religion', *NBER Working Paper* n. 8080.

Saint-Paul, G., (1992), 'Technological choice, financial markets and economic development', *European Economic Review*, **36**, 763–81.

Sala-i-Martin, X. (1990), ‘Lecture notes on economic growth (II): Five prototype models of endogenous growth’, *NBER Working Paper* n. 3564.

Salvadori, N. (1998), ‘A linear multisector model of “endogenous” growth and the problem of capital’, *Metroeconomica*, **49**, 319–35.

Samuelson, P.A. (1939), ‘A synthesis of the principle of acceleration and the multiplier’, *Journal of Political Economy*, **47**, 786–97.

Samuelson, P.A. (1978), ‘The canonical classical model of political economy’, *Journal of Economic Literature*, **16**(4), 1415–34.

Samuelson, P.A. (1985), ‘Modes of thought in economics and biology’, AEA proceedings, *American Economic Review*, **75**(2), 166–72.

Sandilands, R.J. (2000), ‘Perspectives on Allyn Young in theories of endogenous growth’, *Journal of the History of Economic Thought*, **22**, 309–28.

Sawyer, M. (1997), *The NAIRU: A Critical Appraisal*, The Jerome Levy Economics Institute *Working Paper* n. 201.

Scheinkman, J. and M. Woodford (1994), ‘Self-organised critically and economic fluctuations’, *American Economic Review*, **84**, 417–21.

Schmidt-Bleck, F. (1994), *Wieviel Umwelt Braucht der Mensch? MIPS – Das Mass fuer Oekologisches Wirtschaften*, Basel/Berlin: Birkhaeuser.

Schmookler J. (1966), *Invention and Economic Growth*, Cambridge, Massachusetts: Harvard University Press.

Schumpeter, J.A. (1927), ‘The explanation of the business cycle’, *Economica*, **7**(21), 286–311.

Schumpeter, J.A. (1934), *The Theory of Economic Development*, Leipzig, Cambridge, Massachusetts: Harvard University Press.

Schumpeter, J.A. (1939), *Business Cycles. A Theoretical, Historical and Statistical Analysis of Capitalist Process*, New York: McGraw-Hill.

Schumpeter, J.A. (1942), *Capitalism, Socialism and Democracy*, London: Allen and Unwin.

Segerstrom P. S. (1998), ‘Endogenous growth without scale effects’, *American Economic Review*, **88**, 1290–310.

Sen, A.K. (1970), *Growth Economics*, Harmondsworth: Penguin.

Serrano, F. (1995), ‘Long period effective demand and the Sraffian multiplier’, *Contributions to Political Economy*, **14**, 67–90.

Setterfield, M. (1997), ‘“History versus equilibrium” and the theory of economic growth’, *Cambridge Journal of Economics*, **21**(3), 365–78.

Setterfield, M., D.V. Gordon and L. Osberg (1992), ‘Searching for a will o’ the wisp: an empirical study of the NAIRU in Canada’, *European Economic Review*, **36**(1), 119–36.

Shell, K. (1966), ‘Towards a theory of inventive activity and capital accumulation’, *American Economic Review*, **56**, 62–8.

Shell, K. (1967), ‘A model of inventive activity and capital accumulation’, in *Essays on the Theory of Optimal Economic Growth*, Cambridge, Massachusetts: MIT Press, pp. 67–85.

Shell, K. (1974), ‘Inventive activity, industrial organisation and economic growth’, in J. Mirrlees and N. Stern (eds.), *Models of Economic Growth*, London: Macmillan, pp. 77–97.

Shell, K. (1987), ‘Hamiltonians’, in J. Eatwell, M. Milgate and P. Newman (eds), *New Palgrave*, Vol. 2, New York: W.W. Norton, pp. 588–90.

Sheshinski, E. (1967), ‘Optimal accumulation with learning by doing’, in K. Shell (ed.), *Essays on the Theory of Optimal Economic Growth*, Cambridge, Massachusetts: MIT Press, pp. 35–53.

Silverberg, G. and B. Verspagen (1998), ‘Economic growth and economic evolution: a modelling perspective’, in F. Schweitzer and G. Silverberg (eds), *Selbstorganisation. Jahrbuch für Komplexität in den Natur-, Sozial- und Geisteswissenschaften*, Berlin: Duncker & Humblot, pp. 239–64.

Silverberg, G., G. Dosi and L. Orsenigo (1988), ‘Innovation, diversity, and diffusion: a self-organising model’, *Economic Journal*, **98**, 1032–54.

Simon, H.A. (1978), ‘Rationality as process and as product of thought’, *American Economic Review*, **68**, p. 1–16.

Skinner, A.S. and T. Wilson (eds), (1975), *Essays on Adam Smith*, Oxford: Clarendon Press.

Slutsky, E. (1937), ‘The summation of random causes as the source of cyclic process’, *Econometrica*, **5**(2), 105–46.

Smith, A. (1976), *The Theory of Moral Sentiments*, 1st edn 1759, Vol. I of *The Glasgow Edition of the Works and Correspondence of Adam Smith*, edited by A.L. Macfie and D.D. Raphael, In the text quoted as *TMS*.

Smith, A. (1976), *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1st edn 1776, Vol. II of *The Glasgow Edition of the Works and Correspondence of Adam Smith*, edited by R.H.Campbell, A.S. Skinner and W.B. Todd, Oxford: Oxford University Press. Referred to in the text as *WN*, book number, chapter number, section number, paragraph number.

Smulders, S. (1999), ‘Endogenous growth theory and the environment’, in Van den Bergh J. (ed.), *Handbook of Environmental and resource economics*, Chelthenam: Edward Elgar, pp. 656–77.

Snowdon, B. and H.R. Vane (1999), *Conversation with Leading Economists*, Cheltenham: Edward Elgar.

Söderstrom, H. (1976), 'Production and investment under costs of adjustment: a survey', *Zeitschrift für Nationalökonomie*, **36**(4), 369–88.

Soete, L and R. Turner (1984), 'Technological diffusion and the rate of technological change', *Economic Journal*, **94**, 612–23.

Solow, R. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70**(1), 65–94.

Solow, R.M. (1957), 'Technical change and the aggregate production function', *Review of Economics and Statistics*, **39**, 312–20.

Solow, R.M. (1960), 'Investment and technical progress', in K. Arrow *et al*. (eds), *Mathematical Methods in the social sciences*, Stanford: Stanford University Press, pp. 89–104.

Solow, R.M. (1970), *Growth Theory. An Exposition*, Oxford: Oxford University Press.

Solow, R.M. (1974), 'Intergenerational equity and exhaustible resources', *Review of Economic Studies*, **41**, 29–45.

Solow, R.M. (1979), 'Alternative approaches to macroeconomic theory: a partial view', *Canadian Journal of Economics*, **12**(3), 339–54.

Solow, R.M. (1997), 'Is there a core of usable macroeconomics we should all believe in?', *American Economic Review*, **87**(2), 230–2.

Solow, R.M. (1999), 'Neoclassical growth theory', in J.B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, Vol. 1A, Amsterdam: Elsevier, pp. 637–67.

Solow, R.M. (2000a), 'The neoclassical theory of growth and distribution', *Banca Nazionale del Lavoro Quarterly Review*, **53**, 349–81.

Solow, R.M. (2000b), 'Toward a macroeconomics of the medium run', *Journal of Economic Perspectives*, **14**(1),pages.

Sordi, S. (2001), 'Growth cycles when workers save: a reformulation of goodwin's model along Kaldorian–Pasinettian lines', *Central European Journal of Operations Research*, **9**(1–2), 97–117.

Sraffa, P. (1960), *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*, Cambridge: Cambridge University Press.

Stadler, G. (1990), 'Business cycle models with endogenous technology', *American Economic Review*, **80**(4), 763–78.

Steedman, I. (1972), 'The state and the outcome of the Pasinetti process', *Economic Journal*, **82**, 1387–95.

Steindl, J. (1952), *Maturity and Stagnation in American Capitalism*, Oxford: Basil Blackwell.

Stern, D.I. (1997), 'Progress on the environmental Kuznets Curve?', *Environment and Development Economics*, **3**, 175–98.

Stigler, G. (1975), 'Smith's travels on the ship of state' in Skinner and Wilson (1975), pp. 237–46.

Stiglitz, J.E. (1974), 'Growth with exhaustible natural resources: efficient and optimal growth paths', *Review of Economic Studies*, **41**, 123–37.

Stiglitz, J.E. (1986), 'Comments: some retrospective views on growth theory', in D. Diamond (ed.), *Growth, Productivity, Unemployment*, Cambridge, Massachusetts: MIT Press, pp. 50–67.

Stiglitz, J.E. (1993), 'Endogenous growth and cycles', *NBER Working Paper* n. 4286.

Stiglitz, J.E. and A. Weiss (1981), 'Credit rationing in markets with imperfect information', *American Economic Review*, **71**, 393–410.

Stokey N.I. (1995), 'R&D and economic growth', *Review of Economic Studies*, **62**, 469–89.

Stokey, N.L. (1998), 'Are there limits to growth?', *International Economic Review*, **39**, 1–31

Strassman, P. W. (1956), 'Economic growth and income distribution', *Quarterly Journal of Economics*, **70**, 425–40.

Strulik, H. (1997), 'Learning by doing, population pressure, and the theory of demographic transition', *Journal of Population Economics*, **10**, 285–98.

Strulik, H. (1999a), 'Demographic transition, stagnation, and demoeconomic cycles in a model for the less developed economy', *Journal of Macroeconomics*, **21**(2), 397–413.

Strulik, H. (1999b), 'On the mechanics of economic development and non development', Cambridge University DAE-EO-Paper 9904.

Strulik, H. (2000), 'On demographic transition, structural change, and economic growth and stagnation', *Mathematical Population Studies*, **8**, 333–56.

Sturzenegger, F. and M. Tommasi (1994), 'The distribution of political power, the costs of rent-seeking, and economic growth', *Economic Inquiry*, **23**, 236–48.

Swan, T.W. (1956), 'Economic growth and capital accumulation', *Economic Record*, **32**, 334–61.

Tamura, R. (1996), *Human Capital, Specialization, and Industrial Revolution*, Clemson University, *mimeo*.

Targetti, F. (1991), 'Change and continuity in Kaldor's thought on growth and distribution', in Nell and Semmler (1991), 411–20.

Temple, J. and P. Johnson (1998), 'Social capability and economic growth', *Quarterly Journal of Economics*, 113(3), 965–90.

Thirlwall, A.P. (1979), 'The balance of payments constraint as an explanation of international growth rate differences', *BNL Quarterly Review*, n. 128.

Thirlwall, A.P. (1983), 'Foreign trade elasticities in centre-periphery models of growth and development', *BNL Quarterly Review*, n. 146, 249–61.

Thirlwall, A.P. (1986), 'A general model of growth and development on Kaldorian lines', *Oxford Economic Papers*, **38**(2), 199–219.

Thirlwall, A.P. (1987), *Nicholas Kaldor*, Brighton: Wheatsheaf Press.

Thirlwall, A.P. (1991), 'Professor Krugman's 45 degree rule', *Journal of Post Keynesian Economics*, **13**(1), 23–8.

Thirlwall, A.P. (1998), 'The balance of payments and growth: from mercantilism, to Keynes to Harrod and beyond', in G. Rampa, L. Stella and A.P. Thirlwall (eds), *Economic Dynamics, Trade and Growth. Essays on Harrodian Themes*, London: Macmillan Press, pp. 179–204

Thirlwall, A.P. and M.N. Hussain (1982), 'The balance of payments constraint, capital flows and growth rate differences between developing countries', *Oxford Economic Papers*, **34**(3), 498–510.

Toman, M.A., J. Pezzey and J. Krautkraemer (1995), 'Neoclassical economic growth theory and sustainability, in D.W. Bromley (ed.)., *Handbook of Environmental Economics*, Oxford: Blackwell, pp. 139–65.

Townsend, R. (1979) 'Optimal contracts and competitive markets with costly state verification', *Journal of Economic Theory*, **21**(2), 265–93.

Treiman, D. (1977), *Occupational Prestige in Comparative Perspective*, New York: Academic Press.

Trezzini, A. (1995), 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, **14**, 33–66.

Trezzini A. (1998), 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, **19**, 19–32.

Tullock, G. (1967), 'Welfare costs of tariffs, monopolies and theft', *Western Economic Journal*, **5**, 224–32.

U.S. Census Bureau (2001), *Survey of Plant Capacity 1999*, U.S. Department of Commerce, March.

Uzawa, H. (1965), 'Optimum Technical change in an aggregate model of economic growth', *International Economic Review*, **6**, 18–31.

Varri, P. (1990), 'Introduzione', in *Harrod, R.F. Dinamica Economica*, Bologna: Il Mulino, pp. 10–37

Verspagen, B. (2001), 'Economic growth and technological change: an evolutionary interpretation', *STI Working Papers* 2001/1.

Vianello, F. (1985), 'The pace of accumulation', *Political Economy*, **1**(1), 69–87.

Vianello, F. (1989), 'Effective demand and the rate of profit: some thoughts on Marx, Kalecki and Sraffa', in M. Sebastiani (ed.), *Kalecki's Relevance Today*, New York: St. Martin's Press, pp. 164–90.

Vianello, F. (1996), 'Joan Robinson on normal prices (and the normal rate of profits)', in M.C. Marcuzzo, L.L. Pasinetti and A. Roncaglia (eds), *The Economics of Joan Robinson*, London: Routledge, pp. 112–20.

Neumann, J. von (1945), 'A model of general economic equilibrium', *Review of Economic Studies*, **13**, 1–9. English translation of 'Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes', in *Ergebnisse eines mathematischen Kolloquiums*, **8** (1937), 73–83.

Von Waizsaker, C.C. (1974), 'Notes on endogenous growth of productivity', in J. Mirrlees and N. Stern (eds.), *Models of Economic Growth*, London: Macmillan, pp. 101–11.

Wan, H. (1970), *Economic Growth*, New York: Arcourt-Bruce.

Weber, M. (1930), *Protestant Ethic and the Spirit of Capitalism*, trans. T. Parson, New York: Scribner & Sons.

Weiss, Y. and C. Fershtman (1998), 'Social status and economic performance: a survey', *European Economic Review*, **42**, 801–20.

Weitzman, M.L. (1971), 'Shiftable versus non-shiftable capital: a synthesis', *Econometrica*, **39**, 511–29.

Wilson, T. (1976), 'Effective devaluation and inflation', *Oxford Economic Papers*, **28**(1), 1–24.

Winter, S.G. (1971), 'Satisficing, selection and the innovating remnant'. *Quarterly, Journal of Economics*, **85**, 237–61.

Wolfstetter, E. (1982), 'Fiscal Policy and the classical growth cycle', *Zeitschrift für Nationalökonomie*, **42**(4), 375–93.

Woolcock, K. (2000), 'The place of social capital in understanding social and economic outcome', mimeo, The World Bank.

World Commission on Environment and Development (Brundtland Commission) (1987), *Our Common Future*, Oxford: Oxford University Press.

Wray, L.R. (1988), 'The monetary explanation of distribution: a critique of Pivetti', *political economy*, **4**(2), 269–73.

Xie, D. (1994), 'Divergence in economic performance: transitional dynamics with multiple equilibria', *Journal of Economic Theory*, **63**(1), 97–112.

Yang, X. (1999), 'The division of labor, investment and capital', *Metroeconomica*, **50**, 301–24.

Yip, C.K. and J. Zhang (1997), 'A simple endogenous growth model with endogenous fertility: indeterminacy and uniqueness', *Journal of Population Economics*, **10**, 97–110.

Young, A.A. (1928), 'Increasing returns and economic progress', *Economic Journal*, **38**, 527–42.

Young, A.A. (1998), 'Growth without scale effects', *Journal of Political Economy*, **106**, 41–63.

Young, A.A. (1999a), 'Particular expenses and supply curves, (Nicholas Kaldor's notes on Allyn Young's LSE Lectures, 1927–29)', in P.G. Mehrling and R.J. Sandilands (eds), *Money and Growth. Selected Papers of Allyn Abbot Young*, London: Routledge, pp. 391–8.

Young, A.A. (1999b), 'Supply and demand' (from the *Encyclopedia Britannica*, 1928), in P.G. Mehrling and R.J. Sandilands (eds), *Money and Growth. Selected Papers of Allyn Abbot Young*, London: Routledge, pp. 143–6.

Young, W. (1989), *Harrod and his Trade Cycle Group*, London: Macmillan.

Zak, P.L. and S. Knack (2001), 'Trust and growth', *Economic Journal*, **111**, 295–321.

Zenezini, M. (1990), 'Esistono davvero le funzioni neoclassiche di domanda di lavoro?', *Politica Economica*, **8**(April), 19–59.

Zweimuller, J. (2000), 'Schumpterian entrepreneurs meet Engel's law: the impact of inequality on innovation-driven growth', *Journal of Economic Growth*, **5**, 185–206.