The Evolution of the Literature on Technological Change over time: A Survey.

by

Andrea Conte
University of Turin

Number of Pages: 74
The Evolution of the Literature on Technological Change over time: 
A Survey.

Andrea Conte  
University of Turin  

December 15, 2006

Abstract

This paper reviews the emergence and evolution of major topics in economics of innovation. Throughout the paper, particular attention is devoted to the analysis of the cumulative aspects and complementarities between different paths of research over time. Moreover, this survey highlights the crucial relationship between technological change (TC) and economic growth, and the way in which economics literature has dealt with this issue over time. The structure of this survey distinguishes between different decades and it identifies the key debates in the economics literature in each period. Although relevant steps have been made over time, a systematic and satisfactory integration of different theoretical perspectives appears still to be found. In recent years, there have been more sophisticated empirical and theoretical attempts to deal with TC at several, and more disaggregated, levels of analysis. Notwithstanding such advancements, further research is needed to ensure the development of a more general theory of the determinants and the effects of TC. In turn, such theory has to deal primarily with an assessment of both the complementarities between the economic incentives and the internal mechanisms of the so-called “black box” (Rosenberg, 1994), and the heterogeneity which characterises the innovative process of firms across different sectors, countries and over time.

Keywords: Technological change, economic growth, induced innovation, diffusion, evolutionary economics, path dependence.

JEL classification: O30, O40
Contents

1 Introduction 3

2 The Foundation of Economics of Innovation ( - 1960) 5
   2.1 Technological Change among Classical Economists ( - 1910) 5
   2.2 An Analytical Setting for Technological Change (1910 - 1942) 6
   2.3 Technological Change and Growth. A Keynesian Approach (1935 - 1950) 7
   2.4 Technological Change and Growth. A Neo-Classical Approach (1950 - 1960) 8

3 Determinants and Diffusion of Technological Change (1960 - 1970) 9
   3.1 Mono-Casual Explanations of Technological Change 9
   3.2 Technological Change and Capital: Vintage Models 12
   3.3 Technological Change and Labour: Learning Processes 14
   3.4 Technological Change and Diffusion: Epidemic and Probit Models 15

4 Alternative Views on Technological Change (1970 - 1980) 17
   4.1 The Foundation of an Evolutionary Theory of Innovation 17
   4.2 A Two-way Explanation of Technological Change 19

5 Endogenising Technological Change (1980 - 1990) 21
   5.1 Path Dependence 21
   5.2 Technological Change and Diffusion: Some Extensions 22
   5.3 Appropriability and Market Structure 24
   5.4 New Growth Theory in the 1980s 27

6 Recent Approaches to Technological Change (1990 - ) 29
   6.1 New Growth Theory in the 1990s 29
   6.2 The Evolutionary Approach to Technological Change and Growth in Recent Years 33

7 An Agenda for Future Research 35

References 38
1 Introduction

This paper reviews the evolution of the economics literature on technological change (TC). Even narrowing its definition, that is, excluding organizational and managerial change, an exhaustive assessment of the entire research on TC would not be possible in one single survey. The topics and the contributions discussed in each section are therefore very selective. The aim of this paper is twofold. The first objective is to provide a clear overview of the state-of-the-art of the main topics in economics of innovation. They are discussed in those sections related to the years in which the bulk of related literature appeared. Each section of the paper refers to a specific period; the balance between completeness and tractability being solved by providing detailed references to key contributions and other single-topic surveys. The second target is to offer a workable interpretation of the "dynamics" in the literature on TC by emphasising the cumulativeness, the complementarities and the bridging factors between different lines of research.

A very general definition of TC is that "it constitutes certain kinds of knowledge that make it possible to produce (I) a greater volume of output or (2) a qualitatively superior output from a given amount of resources" (Rosenberg, 1982, p. 3). Being the "science which studies human behavior as a relationship between ends and scarce means which have alternative uses" (Robbins, 1932, p. 16), economics has historically dealt with both the allocative problem and, more intensively, the way of "enlarging the cake", namely economic growth. Economists started to care about TC because of its intimate causality link with economic growth. This paper addresses such relationship by drawing explicit links between the evolution of the literature on TC and economic growth.

The structure of this survey is as follows. Section 2 discusses the foundation of the research on TC as it appeared before the 1960s. In particular, the recognition of the importance of TC for economic growth (Section 2.1), Schumpeter's outstanding contribution to the building of the field of economics of innovation (Section 2.2) and the first attempts to build a model of economic growth in both a Keynesian (Section 2.3) and neo-classical framework (Section 2.4). Section 3 discusses the economics debate in the 1960s. These years produced new theories and empirical insights, especially into the interaction between factor prices, product demand / supply, and the rate and direction of TC. Section 4 focuses on the 1970s and on the appearance of evolutionary theory as a viable analytical tool for the analysis of TC. Evolutionary models stressed the concepts of heterogeneity and dynamics of innovative patterns by drawing on Schumpeter's insights into the process of economic development (Section 4.1). Moreover, mono-causal "induced" explanations of TC paved the way to a dynamic two-way representation of the interaction between market forces and innovation (Section 4.2). Section 5 discusses some extensions to previous research conducted during the 1980s. In these years, the development of historically grounded "path dependent" models of innovation (Section 5.1), further insights into the diffusion process (Section 5.2) and the relationship between appropriability, market structure and TC (Section 5.3) were complemented by a new wave of studies on economic growth (Section 5.4). Section 6 describes the recent consolidation of the literature on TC through a large number of theoretical and empirical contributions. This section focuses mainly on the results obtained by new growth theory (Section 6.1) and evolutionary economics (Section 6.2). A descriptive representation of these features, as well as of the structure of this paper, is provided by the following Table 1. Finally, Section 7 concludes this survey by identifying some of the more promising directions for future research.

---

1 Nowadays, a proof of such link is underlined by the common placement of the items "Technological Change" and "Economic Growth" under the same group "O" in the JEL classification system.
TABLE 1. A Descriptive Representation of the Evolution of the Literature on TC over time.

Classical Economists

- Schumpeter

Keynesian Growth Model

- Solow Growth Model

Capital (K): Embodied TC

Labour (L): Learning Processes

Diffusion Models

Mono Casual Explanations of TC

Two-way Explanations of TC

Evolutionary Approach

Appropriability / Market Structure

Growth Theory and Growth Accounting

Spillovers

TC Determinants

Entrepreneurship

New Endogenous Growth Theory

Evolutionary Approach

Path Dependence

Discussion Papers on Entrepreneurship, Growth and Public Policy
2 The Foundation of Economics of Innovation ( - 1960)

This section provides a brief overview of the nascent debate on TC and its relationship with economic growth. This summary starts signalling the recognition among classical economists of TC as a fundamental tool to escape the law of diminishing returns; then, it discusses the building blocks of Schumpeterian approach to growth and TC such as his emphasis on the disequilibrium nature of the innovative process and on the role of entrepreneurs. Finally, this section describes the first formal attempt to model economic growth which opened the way to the growth literature of the 1950s.

2.1 Technological Change among Classical Economists ( - 1910)

The historically new mode of capitalistic production, and the emergence of new institutions and social agents involved in this process, led most classical economists to focus on the process of economic growth, and consequently, on the challenge posed by the law of diminishing returns, which was generally regarded as the main obstacle to a positive growth path of the economy\(^2\). Although a very first general recognition of its importance dates back to Adam Smith (1776)\(^3\), TC, together with the expansion of trade, started to be identified as a crucial factor for weakening the outcome of the law ofdiminishing returns (Ricardo, 1821).

Some decades later, Marx (1867) widened the importance of technological progress in a broader historical perspective by stressing its strong heterogeneity over time and space and the direct relationship between TC and the emergence of capitalistic institutions. Moreover, Marx replaced the concept of stationary state with an alternative historical vision of the economic system characterised by a continuous technology-driven evolution. Indeed, "it is not the articles made, but how they are made, and by what instruments, that enables us to distinguish different economic epochs" (Marx, 1867, p. 180)\(^4\).

The growing attention to the role of TC came to a halt at the end of the 19\(^{th}\) century. In those years, the Marginalist Revolution led by William Stanley Jevons, Carl Menger, Léon Walras and Alfred Marshall rejected the classical labor theory of value by adopting the concept of diminishing marginal utility at the basis of the theory of exchange. This reinforced the concept of stationary state and led to a representation of TC as a mere generator of transient disturbances, being economic change induced by simple adjustments of capital and labour\(^5\). Although the Marginalist Revolution "marginalised" TC in analytical terms, general equilibrium analyses paved the way to the definition and adoption of the concept of production function, namely a set of combinations of technically feasible inputs and outputs, which is still widely adopted in economics literature and, specifically, in many innovation studies (Sections 2.4 and 3.1).

\(^2\)Such challenge, together with a rising trend in population growth and an unsatisfactory growth of means of subsistence in the stationary state, namely a situation in which all the economic variables are fixed relative to each other (Mill, 1879), was at the base of the well-known pessimistic view endorsed by Malthus (1798).

\(^3\)Indeed, he advocated the view that the division of labour - which implies both technological and organizational change - positively affects economic growth.

\(^4\)Marx’s technological determinism is discussed - among others - by Hansen (1921) and Rosenberg (1974a).

\(^5\)The concept of stationary state makes the hypothesis of perfect knowledge more plausibly, if it is coupled to the assumption of perfect competition (Stigler, 1957). This framework understates therefore the role of knowledge in promoting TC, and the degree of heterogeneity among different economic agents (Sections 3.3 and 3.4).
2.2 An Analytical Setting for Technological Change (1910 - 1942)

The first half of the 20th century witnessed the building of a general theoretical framework for the analysis of TC. This was due to the contributions of Hicks (1932) and, in particular, Joseph A. Schumpeter (1883-1950).

The former introduced the "induced innovation" hypothesis, which states that changes in factor prices determine a substitution in favour of the relatively cheaper factor through a factor-biased TC\(^6\). In the context of production function, such substitution is represented by both a movement of the isoquant towards the origin (TC) and a movement along the isoquant (factor substitution)\(^7\). The latter, in his "The Theory of Economic Development" (1934; 1st edn. 1912), set the general taxonomies and definitions which are still widely-used nowadays for both theoretical and empirical research on TC\(^8\). Schumpeter (1934) defined TC as the "carrying out of new combinations of the means of production" and provided his famous "trilogy" by distinguishing the process of TC in invention, innovation and diffusion. The first stage of his taxonomy, the invention process, encompasses the generation of new ideas and it is commonly associated with science and basic research. The second stage, the innovation process, represents the development of new ideas into marketable products and processes\(^9\). This stage is generally associated with technology and applied research and development (R&D) and determines the creation of economic value at a firm level. Schumpeter (1934) distinguished five types of innovation: (1) the introduction of a new good or of a new quality of a good (product innovation), (2) a new method of production (process innovation), (3) the opening of a new market, (4) the discovery of new resources or intermediates, and (5) a new organisational form. Finally, the diffusion phase describes the spread of new products and processes across potential markets. This last component of the "trilogy" allows the measurement of the impact of new technologies on the economy (Sections 3.4 and 5.2).

Besides his analytical classifications, Schumpeter stressed the macro-economic implications of TC, which was seen as the engine which drives economic growth\(^10\). In particular, he strongly advocated the "instability of capitalism" (Schumpeter, 1928) due to the disequilibrium effect of TC on the irregular series of shocks in the economy\(^11\). Exactly this reasoning determined Schumpeter's opposition to the aggregate method of Keynes (1883-1946) and to his demand-based explanation of the business cycle. Indeed, Schumpeter looked at the economic cycle as a by-product of growth\(^12\). In particular, Schumpeter described a process in which entrepreneurs represent the fundamental agents of the process of TC. Indeed, they continually introduce new product or process ("creative destruction") backed by the perspective of extra profits which innovation may guarantee through a temporary market power condition. TC continuously generates new knowledge which, in turn, spills over the whole economy through imitation until it is displaced by subsequent successful innova-

---

\(^{6}\)Hicks (1932) consider an autonomous and an induced component of TC, and defined the latter as the outcome of an entrepreneur's efforts to reduce the use of the relatively expensive factor of production. Therefore, "\textit{a change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind – directing to economising the use of a factor which has become relatively more expensive}" (p. 125).

\(^{7}\)Another view narrowed the concept of TC only to the occurrence of autonomous improvements (Kaldor, 1932; Blaug, 1963) since only in this case the new method of production is not previously known. On the contrary, in the other two possible cases, namely (1) a change in the relative scarcity of factors originating from the supply side or (2) a change in the price of the factors for a given relative scarcity, these new methods are simply not profitable \textit{ex-ante}.

\(^{8}\)An example is given by OECD (2005) definitions and methodologies related to the collection of innovative data.

\(^{9}\)"\textit{Innovation is possible without anything we should identify as invention, and invention does not necessarily induce innovation but produces itself... no economically relevant effect at all}" (p. 84).

\(^{10}\)The importance of TC for long-term economic growth was also underlined by Kuznets (1930).

\(^{11}\)A different perspective, which emphasises the continuity and cumulative nature of TC, was backed by Usher (1954; 1st edn. 1929). Ruttan (1959) discusses a reconciliation of Usher's and Schumpeter's theories.

\(^{12}\)On the contrary, the inverse causality direction was endorsed by Kaldor (1964).
This process directly implies a judgment over the channels of technology diffusion (Sections 3.4 and 5.2) and the most effective market structure to support innovation (Section 5.3). In general, Schumpeter solved the conflict between entrepreneurial activity and perfect competition by sacrificing the latter. He denied that a world of perfect competition ever actually existed and he argued that, even if the conditions for perfect competition could be achieved, this would be undesirable. Perfect competition is irrelevant because it focuses entirely on price competition. Dynamic growth arguments are more important than static allocative efficiency, as well as creative destruction is more important than price competition. In his later contribution, Schumpeter (1942) went even further in his criticism of competitive markets. In particular, he dropped the previously-assessed effectiveness of creative destruction in making monopoly positions vulnerable and he stated an alternative process of TC ("creative accumulation") where innovations are routinised and favoured by monopolistic and oligopolistic positions.

2.3 Technological Change and Growth. A Keynesian Approach (1935 - 1950)

A first formal attempt to model economic growth was independently carried by Harrod (1939) and Domar (1946). The Harrod-Domar model is based on the assumptions of a Leontief aggregate production function and of purely labour-augmenting (Harrod-neutral) technological progress. Moreover, this model assumes a fixed capital-output ratio $\sigma$, savings $S$ represent a fixed proportion of income $S = sY$ and their level equals the level of investment $S = I$. Investment $I$ is a component of the demand for output as well as the determinant of the increase in capital stock (net of capital depreciation). Then $I = \Delta K = \sigma \Delta Y = sY$. More capital accumulation generates higher output and higher income which, in turn, allows higher levels of saving. The Keynesian nature of the model (prices are not flexible) implies the absence of both a guaranteed market-clearing mechanism and a full-employment equilibrium.

The Harrod-Domar growth equation is given by $g = \Delta Y / Y = s / \sigma$.

In this setting, the key to economic growth is to expand the level of investments by encouraging saving and/or lowering the capital-output ratio $\sigma$, for instance, by generating technological advances which increase the productivity of the investment and, therefore, enable firms to produce more output with less capital.

Although several shortcomings characterise the Harrod-Domar setting, this model has been considered as the precursor of (exogenous) growth theory developed during the 1950s.

\[13\] The concept of creative destruction was criticised on the ground of some empirical evidence which shows a peacefully coexistence of old and new technologies (Strassman, 1959).

\[14\] A Leontief aggregate production function takes the following form: $Y = \min\{x_1, x_2\}$ where $Y$ is the quantity of output and $x_1$ and $x_2$ are quantities of inputs (Leontief, 1956; Diewert, 1971). Labour-augmenting TC implies that with the same amount of capital one needs less and less labour to produce the same amount of output.

\[15\] First, the law of diminishing returns would suggest a flexible capital-output ratio $\sigma$. Second, keeping the assumption of a fixed relative price of labour ($L$) and capital ($K$) implies that growth is enough to maintain full employment. Third, although it is possible to obtain a steady-state equilibrium growth path, the Harrod problem is how an economy reaches and maintains this path. Finally, the idea that investment is only influenced by output is criticised by Barro and Sala-I-Martin (1995) based on the non-optimality of a saving behaviour at a constant rate for utility-maximising agents when the marginal product of capital is zero.
2.4 Technological Change and Growth. A Neo-Classical Approach (1950 - 1960)

At the beginning of the 1950s, the prevailing view about economic growth emphasised the importance of an increasing capital-labour ratio $(K/L)$. This idea was challenged during this decade by an alternative approach which stressed the disembodied nature of TC, which allows, for some given input vectors, the production of output vectors not previously feasible. In this setting, TC was depicted by movements over time of the production possibility frontier and, as a result, it was generally represented as a problem of maximisation under constraints where its rate and direction derived from the rational choice of the representative firm (Swan, 1956; Samuelson, 1958).

Solow (1956) provided a theoretical model of long-run growth where a purely labour-augmenting TC is added to the Cobb-Douglas aggregate production function $Y = F(K, LA)$. This is characterized by constant returns to scale and diminishing marginal returns to capital $(K)$ and labour $(L)$\textsuperscript{16}. The latter allow to relax the Harrod-Domar assumption of a constant capital-output ratio $\sigma$ and to introduce flexibility between factors of production. While the savings function adopted, that is $S = sY = I$, is the same as in the Harrod-Domar model, a further difference between the two frameworks refers to their dynamics, driven by capital in the Harrod-Domar model and capital per worker $(k)$ accumulation in the Solow’s (1956) one. By expressing the variables in unit-of-effective-worker $(LA)$ values, this follows: $y = \frac{Y}{LA}; \quad k = \frac{K}{LA}; \quad y = f(k)$. In this setting, $LA$ grows at $n + a$, where $n$ and $a$ represents the increase in population $L$ and technology $A$ respectively. Since $k$ is constant, both $K$ and $Y$ must also be growing proportionally at $n + a$. Therefore, $\frac{Y}{L}$ grows at $a$, which indicates the role of TC as the only determinant of economic growth in the long-run.

Following Fabricant’s (1954) view that the growth of conventional inputs does not explain much of the observed growth in output, Solow (1957) decomposed different sources of per-capita output growth in order to quantify the influence of $L$ and $K$ accumulation. TC was considered Hicks-neutral and fully disembodied, being proxied by a time variable $t$ in the aggregate production function $Y = F(K, L, t)$\textsuperscript{17}. As a result, he obtained evidence of the marginal role of $L$ and $K$ accumulation while a so-called ”Solow residual”, labelled TC and exogenous to the model, accounted for 87.5% of output growth\textsuperscript{18}.

Solow’s (1957) model has originated several productive streams of research in the following decades. These have focused on the relaxation of some assumptions of the original setting, such as the neutrality of TC and its explanations (Section 3) or the hypotheses of perfect competition and constant returns to scale (Sections 5.3, 5.4 and 6.1). Moreover, the usefulness of the concept of aggregate production function was at the hard core of the Cambridge-Cambridge controversy in the theory of capital during the 1950s\textsuperscript{19}.

\textsuperscript{16}Moreover, the concept of production function relies on the assumptions of separability and aggregation. It is generally assumed continuous and continuously differentiable.

\textsuperscript{17}Hicks-neutrality means that the ratio of capital and labour marginal products remains constant. In the Cobb-Douglas case, this is similar to Harrod-neutrality, that is $Y = AF(K, L) = F(K, LA)$. Steedman (1985) discusses the different definitions of neutrality and some related conditions for their plausibility.

\textsuperscript{18}Economics literature has provided several methods to express the residual (Domar, 1961). Although it was associated with TC, ”the indicated importance of this element may be taken to be some sort of measure of our ignorance” (Abramovitz, 1956, p. 11). Indeed, other inputs than pure TC, such as “slowdowns, speed-ups, improvements in the education of the labor force, and all sort of things” (Solow, 1957, p. 402), including also measurement errors (Kendrick, 1956; Jorgenson and Griliches, 1967), may affect productivity increase.

\textsuperscript{19}This controversy focused on both the measurement problems related to the aggregation of capital (Sraffa, 1960) and the possibility of reswitching, namely a situation where a production technique is best cost-minimising at high or low rates of profit, but it is not at an intermediate rate of profit (Samuelson, 1966b and Samuelson and Modigliani, 1966). Critics of the concept of aggregate production function (Robinson, 1955 and Kaldor, 1955 and 1957) argued that only a change in the rate of saving, rather than changes in the rate of real profit or wage, would affect the capital coefficient and, therefore, the capital intensity $\frac{K}{L}$ and the (flexible) capital-output ratio $\frac{K}{Y}$. In order to characterize different forms of TC, they adopted alternative techniques such as the identification of wage-profit frontiers (Harcourt,
3 Determinants and Diffusion of Technological Change (1960 - 1970)

TC entered the economics debate of the 1950s as the crucial unexplained factor affecting economic growth. In the successive decade, the need of a more accurate growth theory pushed theoretical and empirical research into the nature and determinants of TC. Since the most important sources of growth in modern society are "the stock of technological knowledge embodied in tangible records and in the personal skills and habits of the population" (Kuznets, 1947, p. 12), research increasingly focused on "the embodying of technical change in capital inputs... and on the embodying of ostensibly superior technical knowledge and skill in the labour force through the agency of education" (David and Van de Klundert, 1965, p. 357). In turn, this implied a major research focus on both (1) labor- or capital-augmenting TC and (2) unmeasured quality improvement in the inputs. The former recalls the debate on the direction of TC and its relationship with economic forces (Section 3.1) while the latter refers to two different streams of literature assessing the "quality improvements" of the two production factors: capital (Section 3.2) and labour (Section 3.3). Finally, another line of research emerged during the 1960s and discussed the diffusion process of new technologies, namely the third stage of Schumpeterian trilogy. These studies originated from the importance of diffusion in generating the link between TC and economic growth, since the productivity-enhancing effect of superior technologies depends upon their utilization in the appropriate places (Section 3.4).

3.1 Mono-Casual Explanations of Technological Change

Economics literature has provided different approaches for assessing the relationship between TC and factor prices and, consequently, the appearance of particular types of innovation in the economy.

Hick's (1932) statement was widely discussed during the 1960s. Salter's (1960) rejection of Hicksian argument was based on the correct characterization of an entrepreneur’s actual target. In addition, Salter (1960) underlined that the inducement to innovation is lacking in a competitive equilibrium framework, since this implies the equalization of marginal products and factor prices, which causes all factors being equally expensive to firms. This means that a representative firm considers each factor-price ratio as a parameter and, thus, factor-biased innovations can not be interpreted as the collective response to a change in the factor-price ratio (Elster, 1983, p. 102).

Notwithstanding Salter’s (1960) arguments, the theory of “induced innovation” was developed both at a macro and a micro economic level and gained large support. Its macro-economic version emerged since neoclassical analyses based simply on factor substitution appeared unable to explain (1) the large differential in the growth rates of $K$ and $L$ and (2) the factor share constancy witnessed during the 20th century, despite rapidly rising wage rates and the large substitution of $K$ for $L$.

Kennedy’s theory (1964) was based on the concept of “innovation possibilities function” (IPF) rather than on a static neo-classical production function. He combined the IPF with an entrepreneur’s maximisation behaviour which seeks to maximise, subject to the frontier, the current

---

20In particular, "the entrepreneur is interested in reducing costs in total... when labor costs rise any advance that reduces total cost is welcome, and whether this is achieved by saving labor or saving capital is irrelevant" (p. 43).


22The concept of IPF is similar to Kaldor’s (1961) technical progress function (TPF), which describes a linear relationship between the accumulation of capital per worker and productivity growth, and a non-linear relationship between the latter and TC. A similar model is provided by Von Weizsacker (1966).
rate of cost reduction and discussed the effect of changes in relative factor shares on bias in invention. Therefore, he defined \( \lambda = \frac{L}{C+L} \) and \( \gamma = \frac{C}{C+L} \), where \( \lambda \) is the share of labour cost in total cost and \( \gamma \) is the share of capital cost in total cost. With constant factor prices, this relationship holds \( r = \lambda p + \gamma q \) where \( r \) represents the proportional reduction in unit costs due to the introduction of an innovation; \( p \) and \( q \) represent, respectively, the amount of labour and capital reduction for a given product unit. This means that an entrepreneur’s choice of a cost-reducing innovation depends on economic weights \( \lambda \) and \( \gamma \) and, thus, can not be considered independently given, as sustained by Salter (1960). The IPF can be obtained by maximising \( r \) under the constraint \( p = f(q) \), which leads to \( \frac{dp}{dq} = -\frac{\gamma}{\lambda} \). The equilibrium values of the weights are therefore determined by the bias in innovation possibilities \( \left( \frac{dp}{dq} \right) \) when \( p = q \).

However, Kennedy (1964) did not capture any causality relationship between the accumulation of TC and the trade-off between labour and capital-augmenting TC since the shape of the IPF is independent from the factor bias of TC. Moreover, his setting lacks any clear explanation of how R&D activities are financed and priced.

Two extensions of a macro-economic theory of "induced innovation" were discussed by Drandakis and Phelps (1966) and Samuelson (1965). The former provided a model where the IPF indicates the maximum rate of labour-augmentation corresponding to a given rate of capital-augmentation. The latter replaced Kennedy’s (1964) hypothesis of a constant rate of interest with a path of relative capital-labour accumulation and, thus, derived a steady-state solution in case of induced labor-augmenting inventions.

Although the growth-theoretic version of the induced innovation model developed in the 1960s has been unproductive of empirical research and is no longer viewed as an important contribution to growth theory (Ruttan, 1997, p. 1521), it helped neoclassical growth theory to define the existence of an equilibrium growth path which is consistent only with labor-augmenting (Harrod-neutral) TC (Uzawa, 1961a; Amano, 1964). Its lack of micro-foundations, namely a description of the way in which the entrepreneur finds the frontier and moves along it, pushed research towards the definition of a solid micro-founded theory of induced innovation.

Ahmad (1966) related the "induced innovation" hypothesis to a firm’s profit-maximizing R&D decision within the framework of the traditional comparative static analysis. His study is based on the conceptualization of the (historical) innovation possibility curve (IPC), not only of Cobb-Douglas type as Kennedy’s IPF (1964). The IPC is the envelope of all the alternative isoquants which represent a given output on different production functions with narrow possibilities for substitution at any given time. This is not the result of any economic choice, it is a purely technological or laboratory question. The economic consideration would come into play in choosing a particular isoquant (representing a particular production function) out of various isoquants belonging to a particular IPC (p. 347). The convexity of the IPC ensures that TC would respond to a relatively

---

23 Kennedy (1964) neglected the importance of changes in relative factor-prices since they are not essential for a theory of induced bias in innovation. There is a good deal to be gained by presenting the theory in the first instance in a model in which relative factor prices do not change (p. 542).

24 A labour-saving technological bias is such that \( -\frac{dp}{dq} > 1 \).

25 Indeed, "the true case of induced innovation requires at least two productive activities; production and invention. If there is no invention then the theory of induced innovation is just a disguised case of growth theory with exogenous technological change" (Nordhaus, 1973, p. 210). Exactly such extension to a two-sector modelling strategy provided the basis of new growth theory in the 1980s (Section 5.4).

26 An ancillary outcome of this model was the rejoinder between the two authors (Kennedy, 1966; Samuelson, 1966a).

27 This link has been discussed also by Binswanger (1974).

28 Ahmad (1966) criticised also Kennedy’s (1964) hypothesis that the factor share saved by an invention is independent of the amount of that factor used. This started a further debate traceable on The Economic Journal in those years (Ahmad, 1967a and 1967b; Fellner, 1967; Kennedy, 1967).
higher factor price by using less of that factor. The IPC is also neutral, which allows to isolate the effect of changes in factor prices on the nature of TC and implies that a TC responding to any given relative factor price at time \( n \) is neutral to previous TC, which responded to the same relative factor price at time \( n - 1 \).

Although this theory appealingly attempted to describe a micro-economic process in which both economic forces and technology have a role, the sensitivity of the IPC specification to any parameter change weakened such setting. However, its major shortcoming reflected the inability of the traditional comparative static framework to provide a theory of induced invention (Mansfield, 1968). Indeed, although two factors, a learning process (Fellner, 1961) and some degree of market power (Kamien and Schwartz, 1968), may lead to the analysis of "induced innovation" in a comparative static framework, this micro-level theory is still driven by exogenous changes in a firm’s economic environment while its internal innovative features play no role in this setting.

A related stream of research has emphasized the relative importance of market demand in inducing TC, which is depicted as a pure response to profit opportunities (Kamien and Schwartz, 1982). Bloom (1951) provided evidence of a stronger effect on TC of other market forces, such as demand evolution, rather than cost reducing behaviours. In his study of the invention and diffusion of hybrid maize, Griliches (1957) stated that "the process of innovation, the process of adapting and distributing a particular invention to different markets and its acceptance by entrepreneurs, is amenable to economic analysis" (p. 522). However, it was Schmookler (1966, p. 136) that offered explicitly the view that TC is driven by demand evolution by providing evidence of positive correlation between investments of user-sectors and patents of capital-good sectors in four industries (railroads, agricultural machinery, paper, petroleum). Since the former represent the demand for the capital-good sectors, delays of patenting activity support the hypothesis that changes in the patterns of demand are a more effective determinant of TC than advancements in basic science.

On the contrary, the "scarcity push", or "technology push", hypothesis stressed the role of the internal features of TC in its evolution, which is induced by the relative scarcity of conventional inputs, namely capital, land and labour. "Scarcity push" and "induced bias" can be easily interrelated because of the existing relationship between a relative factor scarcity and a higher price of the same factor (Kaldor, 1932). The "scarcity push" version of "induced innovation" theory dates back to Marx’s (1867) assessment of the causality relationship between labour scarcity, the exhaustion of the reserve army of labor, and the substitution of machinery for labor. Although based on Rotherbarth (1946), this thesis was explicitly discussed by Habakkuk (1962), which argued that the higher ratio of land to labour raised real wages in American agriculture more than in the UK during the 19th Century. Such process resulted in both a quantitative change, the substitution of capital for labor, and a qualitative change, the introduction of capital embodying labour-biased TC (James and Skinner, 1985). This hypothesis, and especially its quantitative part, was addressed by Temin (1966) and Fogel (1967) which argued that the scarcity of labour was only one of the explanations, although the most amplified one, for describing Anglo-American differences.

The "technology push" version was initially discussed during the 1960s and fully elaborated in the following decade (Section 4.2). Its origin lay in the recognition of the importance of basic research (Nelson, 1961) as a key tool for promoting significant advances in knowledge which "are often not directly and immediately applicable to the solutions of practical problems" (Nelson, 1959, 1962).

---

29Such "demand induced" model received further support by Vernon (1966), Lucas (1967), Ben-Zion and Ruttan (1975), Kleinhecht and Verspagen (1990).

30Indeed, Temin (1966) stressed the importance of the education of the workforce (Section 3.3) and the extent of the American market. The market size effect has been recently discussed by Acemoglu (2002a). He advocated the view that the market size effect leads to TC which favours abundant factors whereas the price effect pushes TC directed at scarce factors.
p. 302). This line of research emphasised the effect of external economies and the "success breeds success" tendency in creating knowledge that "cannot be costlessly and timelessly transferred among firms" (Phillips, 1966, p. 304). In turn, this called for the recognition of the importance of scientific personnel and the closeness of science and industrial technology (Musson and Robinson, 1960). The central argument is that TC can not simply be narrowed to demand considerations. Indeed, while the amount of a firm’s technological investment depends on the expected profits, the amount of resources devoted by a government depends on other social and political factors such as the closeness of an industry to defence, public health and other social needs (Mansfield, 1968).

The theoretical setting described in this section has provided a first assessment of the interaction between economic factors and TC. However, a comprehensive theory of the relationship between TC and growth requires a wider approach to the determinants of TC evolution, which includes an assessment of the innovative role played by input factors, as well as a description of TC diffusion across the economy. This analysis starts in the next section by discussing the effect of new capital goods as a major channel of technology adoption.

3.2 Technological Change and Capital: Vintage Models

One of the major shortcomings of Solow’s growth model (1957) refers to the assumption of a fully disembodied TC in the context of an aggregate production function. Such hypothesis implies that TC is unrelated to neither the investment variable nor the composition of the labour force.

During the 1960s, a growing unsatisfaction arose from the separate treatment of TC and capital. Indeed, this meant neglecting that "many if not most innovations need to be embodied in new kinds of durable equipment before they can be made effective" (Solow, 1960, p. 91). Moreover, capital itself requires a complex analysis since it "is usually not hired but purchased, it lasts long, and its cost is ambiguous" (Domar, 1962, p. 602). These aspects generated a debate which focused on both the correct measurement of the embodied-knowledge component of capital and the contribution of the whole capital stock to productivity growth (Abramovitz, 1962; Solow, 1962a and Denison, 1964).

A line of research, that is vintage models of TC, emerged in these years and discussed the effect of TC embodied in different vintages of capital equipment on a firm’s productivity. Two common features of these models were (1) the assumption of technological homogeneity within each vintage of capital and technological heterogeneity between different vintages and (2) the hypothesis of exogeneity of the invention process in the capital-good industry.

The embodiment of TC in new capital equipment provided several analytical advantages. First, by differentiating capital stock between machines of different vintages, these models were able to recognise heterogeneous behaviours among firms and led to a more realistic formulation of the investment decision rule followed by the firms. Indeed, such choice depends on a firm’s adoption costs and its expected increase in productivity; firms are therefore expected to make different optimal decisions regarding when to invest in new machines which embody better technology. In turn, this provides a more reasonable interpretation of heterogeneous behaviours rather than the mere occurrence of sub-frontier behaviours (Aigner et al., 1977 and Fare et al., 1994). Second, in a vintage framework, multi-sector models can depict technology transfers among sectors and deal with the optimal intersectoral allocation of inventive effort (Drandakis and Phelps, 1966). Third, vintage models allow for a distinction between embodied and disembodied TC, where the former does not change the productivity of existing capital while the latter affects all vintages in the same way and it is not necessarily related to a firm’s investment (Solow, 1962b). Fourth, vintage models make possible the consideration of both the timing and the rate of TC, since firms change...

---

31 A recent example is given by Parente (1994) which discusses the choice of introducing a new technology.
their capital composition by lengthening (the operating life) and deepening (the multiplication of machines without any change of their longevity) their own capital stocks. Indeed, an acceleration of the pace of capital accumulation, by reducing the age of the capital stock, speeds the rate at which embodied TC is incorporated into production (Nelson, 1964). Fifth, vintage models provide an explanation of two major empirical facts, namely the growth differentials between different economies and the decline of the capital-consumption goods price ratio (Greenwood et al., 1997). Finally, these models allow to consider the role of monetary variables when assessing the relationship between TC and capital investment.

Capital vintage models can be distinguished by looking at the assumptions related to the substitutability between capital and labour, namely a change in the capital-labour ratio, and its timing.

On one side of the spectrum, there are "putty-putty" vintage models where capital has a flexible attribute so that both ex-ante and ex-post substitutability between capital and labour is allowed. Solow’s vintage model (1960) belongs to this category. He replaced the hypothesis of capital homogeneity from his previous 1956 model with a quality-adjusted index number of the amount of capital and related TC to a firm’s investment in new vintages of capital. This new input measure appeared in his modified Cobb-Douglas aggregate production function, where the other input was (homogeneous) labour. In this setting, a purely exponential-in-time capital-augmenting embodied TC improves the productivity of new capital goods while old vintages are unaffected. The two models (1956 and 1960) are therefore equivalent for the (trivial) case of absence of TC.

"Putty-clay" vintage models lie in an intermediate position since only new capital is considered "putty". The notion that labor can be combined with new investment in variable proportions and with existing capital only in fixed proportions was introduced by Johansen (1959). The difference within this class of models refers to the treatment of capital longevity and its labor intensity. For instance, an optimal capital replacement period is found dependent on the rates of TC and investment (Massell, 1962), on the interest rate and the anticipated course of wages (Phelps, 1963). This setting introduces a new dimension to the relationship between investment and productivity growth. Indeed, "the rate of replacement investment is a function of the relative prices of labour and real investment. When real wages are high, standards of obsolescence are high and a high level of replacement investment ensures rapid adjustment to new methods" (Salter, 1960, p. 73).

On the other side of the spectrum, "clay-clay" vintage models do not allow neither ex-ante nor ex-post substitution between capital and labour. Firms can modify the average capital-labour ratio only through changes in the lifetime of equipment. Kaldor and Mirrlees (1962) discussed a "clay-clay" model by adopting the concept of technical progress function (TPF) whereas Solow et al. (1966) assumed a fixed-coefficient technology with embodied technical progress in the context of an aggregate production function.

Although the embodiment of TC in new capital goods reflects the crucial role of capital-good industries in promoting invention and diffusion of TC in the economy (Rosenberg, 1963a and 1963b), this cannot capture the innovative contribution obtained by changes in practices which are implemented with existing equipment. Next section discusses the main features of this type of TC.

32 However, it is possible to revert this relationship by assessing the impact of TC on the rate of capital accumulation. Indeed, embodied TC increases the marginal productivity of capital more than that of labour. In turn, this increases the relative demand for capital and, thus, the growth rate of capital-labor ratio (Abramovitz and David, 1973).

33 Monetary factors affect the degree of capital intensity of an economy (Tobin, 1965). TC can be seen as the result of the allocation of resources away from current production into "technological investment". In turn, this implies that the rate of interest is negatively related to the rate of TC (Lucas, 1967).
3.3 Technological Change and Labour: Learning Processes

The first attempt to develop a knowledge-based theory of endogenous growth was started during the 1960s. Vintage models stressed technological differences among capital input (Section 3.2). At the same time, other scholars widened this analysis by advocating the formal symmetry between improvements in the quality of both the capital stock and the labour force. Since shifts in technology can be divided into an embodied and a disembodied component, the task became the representation of the latter, namely a situation where more output is obtained from unchanged inputs, with no investment required. Arrow (1962a) provided a model which ascribed TC to experience, namely the "activity of production which gives rise to problems for which favorable responses are selected over time" (p. 156). Experience implies the rationalization of the production process and an increase in workers' expertise through the acquisition of knowledge, or learning. Although he recognised that the accumulation of learning depends also on the existence of the so-called "Horndal" effects (Lundberg, 1959) and the role played by institutions, such as schools and universities (Nelson and Phelps, 1966), Arrow (1962a) narrowed the concept of learning to a by-product of ordinary production (proxied by a cumulative gross investment variable). Learning takes place therefore only in the capital-good industry and it is characterised by sharply diminishing returns. As well as Solow (1960), Arrow (1962a) embodied TC in new capital goods. However, he adopted a TPF characterised by Leontief fixed coefficients instead of a Cobb-Douglas production function. In this setting, learning is a pure public good, namely non-rival and non-excludable. This implies that the incentive to innovate is not compensated by the market because of the presence of externalities. Such assumption, together with the absence of invention costs, results in a model characterised by constant returns to scale at the firm level and increasing returns to scale at the aggregate level while it still maintains a competitive equilibrium framework (labour and capital are paid their marginal products). Increasing returns to scale are a crucial determinant of productivity growth (Young, 1928). Moreover, learning as a productivity-increasing mechanism, together with increasing returns to adoption due to positive adoption externalities (Berndt et al., 2000), characterises market penetration and diffusion of many technologies (Tsur et al., 1990). Finally, since learning is proxied by cumulative gross investment of all firms, this generates a scale effect. The model of Arrow (1962a) represents the starting point of new growth theory (Section 5.4).

Two extensions to his concept of learning by doing have been provided some years later. Learning by using represents the demand-side counterpart of learning by doing (Rosenberg, 1982). While the latter describes a form of learning which takes place at the manufacturing stage, high degree of system and technological complexity affects the value of a product for consumers as they gain experience using it. Finally, a third form of learning, learn to learn, reflects the positive relationship between the stock of knowledge and the ability to process information (Stiglitz, 1987).

---

34 Arrow's (1962a) concept of learning by doing, also labelled as “progress function”, “learning curve” or “experience curve”, is related to the analyses of Hirsch (1956) and Verdoorn (1949 and 1956). For a survey of the literature on experience curves, see Day and Montgomery (1983), Dutton and Thomas (1984) and Argote et al. (1990).

35 Non excludability (a non-payer agent cannot be excluded from the benefits of the good) and non rivalness (the marginal cost of an additional agent consuming the good, once it has been produced, is zero) allow to distinguish between private and public goods and, within the latter, between pure public goods, public goods and common pool of resources. The theory of public goods has been originally proposed by Samuelson (1954). Scherer (1967b) explored the implication of rivalry for a firm’s investment in R&D.

36 Nevertheless, his model implied that the growth rate of output was limited by the growth rate of the labour supply, being labour productivity an increasing and concave function of knowledge (Romer, 1986). A specific case in which Arrow (1962a) setting determines an endogenous growth path is discussed by D’Autume and Michel (1993).

37 Moreover, his concept of learning by doing has also been widely applied to the study of market conduct and performance within a decision theoretic approach (Spence, 1981; Fudenberg and Tirole, 1983).
3.4 Technological Change and Diffusion: Epidemic and Probit Models

The third stage of the Schumpeterian trilogy matters since TC has to be exploited and diffused through licensing, imitation or simple adoption to generate productivity growth. Indeed, a diffusion process which goes on too slowly affects negatively the degree of social benefits obtained from TC.

An analytical advantage of the distinction between embodied and disembodied TC is the possibility to identify the channels of technology diffusion, such as the acquisition of capital goods. As well as consumers, firms evaluate the attributes of capital goods, i.e. their different input ratios, and choose consequently (Lancaster, 1971). Learning by doing affects the diffusion process as well. Indeed, the externality associated to the accumulation of production experience represents an important determinant of technology adoption, since it introduces dynamic increasing returns to scale for individual technologies.

Different lines of research have addressed the determinants and patterns of technology diffusion. Economic historians have stressed the impact of geographic movement of skilled workers on the diffusion of technologies (Scoville, 1951; Landes, 1969). Other institutional factors, such as the presence of business trade association or the use of a common capital-good supplier, have contributed to the diffusion process as well by lowering the cost of acquiring information (Graham, 1956; North, 1958 and 1968; Knauerhase, 1968; Walton, 1970; Saxonhouse, 1974).

Economic theory has provided many useful insights into the analysis of technology diffusion. By borrowing the terminology from the analyses of diseases, epidemic theories of diffusion (Bain, 1962; Bass, 1969) have emphasised the positive externality obtained by the transfer of information. Technology spreads through contacts between actual users of a new technology and potential adopters over time. This process is represented by the following differential equation \( \frac{\partial s}{\partial t} = \beta s(1 - s) \)

where \( s \) represents the share of users which have already adopted the new technology and \( \beta \) describes the “contagiousness” of the disease, namely the speed of the diffusion process. One of the main empirical regularities found by these studies is that the diffusion of new superior technologies is a gradual rather than instantaneous process (Jaffe and Stavins, 1994). Diffusion is characterised over time by a sigmoid-shaped curve, which is given by the logistic function obtained by this equation. This means that the probability of a non-user to become user increases with the growing popularity of the technology. Applied epidemic studies show that \( \beta \) depends on economic forces. Expected profits strongly affect both the timing and the rate of technology adoption. In particular, the more profitable the new technology is, the steeper the diffusion curve appears (Griliches, 1957). The probability of technology adoption is also positively affected by the proportion of firms already using it (which confirmed the validity of the main assumption behind epidemic models), firm size (David, 1966), the age of its equipment and its financial liquidity (Mansfield, 1963b) whereas it is negatively related to the size of the investment required for its installation (Mansfield, 1961).

The origin of the information distinguishes epidemic models between those with "internal" influence and those with "external" influence. The former assume that there are some initial users. If nobody has adopted the technology, there is no information to transfer and nobody will learn about the technology. On the contrary, information in the latter is spread from an external source, such as the government or technology suppliers. A further distinction is between intra-firm and inter-firm diffusion. Intra-firm diffusion can be defined as the study of the time path of \( Z_{ijt} = X_{ijt}/Y_{ijt} \)

where \( Y_{ijt} \) is the total output produced by firm \( i \) (\( i = 1...n_{jt} \)) in industry \( j \) in time \( t \) whereas \( X_{ijt} \) is the output produced by adopting the new technology (Karshenas and Stoneman, 1995). In contrast, theories of inter-firm diffusion aim at explaining the time it takes for different firms facing different

---

38 Education displays a positive effect on the speed of technology diffusion (Nelson and Phelps, 1966).

39 Lekvall and Wahlbin (1973) provide an epidemic model in which both internal and external influences are present.
market conditions to adopt a new technology.\textsuperscript{40}

Several shortcomings characterise epidemic models. First, technology has been generally considered constant over time (Dixon, 1980 and Griliches, 1980).\textsuperscript{41} Second, these models have ignored a possible change in the profitability of adoption along the diffusion path. Indeed, diffusion occurs gradually even though the new technology results profitable for all firms. Third, the disequilibrium effect of diffusion calls for a revision of the conceptual basis for measuring diffusion rates (Gold, 1981). Fourth, the underlying theory on the role of learning, that is the acquisition of information, is unsatisfactory since firms are considered as passive imitative recipients rather than active seekers of information.\textsuperscript{42} Finally, these models have generally assumed a homogeneous population of potential users. They can, therefore, explain which innovation diffuses fastest but not which firms become actual users of the technology first since they are not based on decision-theoretical foundations.

Probit models (also called threshold-value or rank models) of technology diffusion relaxed this last assumption and emphasized technology adoption as the result of a value-maximizing decision made by a heterogeneous population of potential adopters (Bonus, 1973; Sommers, 1980). At any moment in time, a threshold point splits potential users between adopters and not adopters since adoption is costly and associated to heterogeneous returns to its users. Probit models obtain the typical sigmoid diffusion path by assuming a single-peaked (normal-like) distribution of potential users and gradual movements to the right of this threshold due to simultaneous price decreases and quality improvements of the technology. Both firm-related and technology characteristics affect adoption choice, rather than a firm’s mere information about technology existence.

Studies on technology diffusion have borrowed some insights from the "induced innovation" literature. For instance, David (1969) has stressed the role of an increasing wage rate in lowering the threshold over time and favouring the diffusion process. In particular, firms compare adoption costs (additional capital costs) and gains (labour cost savings) induced by TC and choose consequently. Davies (1979) has related a firm’s adoption choice to an expected measure of profitability, defined as pay-back period, and several firm-level and technology characteristics. If the value of the payback period is above an acceptable-to-the-firm threshold, the new technology will be adopted. This threshold moves to the right over time since quality improvements and risk reduction facilitate technology adoption.

Several shortcomings are common to both epidemic and probit models. First, they narrow the analysis to the only consideration of demand-side factors, so neglecting the role that technology suppliers play in the diffusion process.\textsuperscript{43} Second, they do not adequately incorporate uncertainty, expectations and risk aversion. Finally, these models neglect any strategic component in the adoption choice since they assume that the benefits obtained from adoption are unrelated to the number of firms already using the new technology.\textsuperscript{44} A new line of research in the 1980s has dealt exactly with these issues; this new wave of diffusion studies is discussed in Section 5.2.

\textsuperscript{40}The way in which information enters into the analysis provides a distinction between inter-firm and intra-firm diffusion models. In the former, information is external to the firm while in the latter firms find their source of information internally. Even though analytically different, some evidence suggest that "there is a considerable amount of unity and similarity between the two diffusion processes" (Mansfield, 1963a, p. 358). Several extensions to epidemic models have been developed over time (Mansfield, 1989).

\textsuperscript{41}Chow (1967) discussed quality improvements on the demand and diffusion of computers.

\textsuperscript{42}Glaister (1974) provided a demand-based model of technology diffusion which also included the cost of a firm’s technology adoption and advertising.

\textsuperscript{43}Demand-side factors determine the parameter reflecting the “contagiousness” of the disease in the epidemic models and the location of the adoption threshold in the probit model.

\textsuperscript{44}The positive externality generated by technology adoption in an epidemic setting suggests that a socially suboptimal diffusion path may happen whereas probit models do not imply a negative evaluation of slow diffusion.
4 Alternative Views on Technological Change (1970 - 1980)

The 1970s witnessed the coexistence of a marked diffusion of TC and a worldwide decrease of the rate of productivity growth which occurred roughly in line with the downswing of the fourth Kondratiev long wave (Petit, 1995). This led economic research to focus more on short-term issues, such as the real business cycle and the effect of rising energy prices (Berndt and Wood, 1986), while the interest for growth theory declined. Scholars turned to discuss the importance of TC per se. Evolutionary economics put a strong emphasis on micro-foundations by building a heterodox behavioural theory of firms. At the macro level, it applied Schumpeterian concepts of disequilibrium and dynamics to the representation of the economic system (Section 4.1). The 1970s witnessed also a growing critique of mono-causal explanations of the evolution and direction of TC. As a result, the ”demand pull” theory paved the way to the recognition of an interactive relationship between TC and economic factors over time (Section 4.2).

4.1 The Foundation of an Evolutionary Theory of Innovation

Several contributions by Nelson and Winter (1973; 1974; 1975; 1977), up to ”An Evolutionary Theory of Economic Change” in 1982, faced the lack of evolutionary foundations in economics, as underlined by Veblen (1898), and constituted the body of a modern evolutionary theory of TC.

Evolutionary research applies the biological principles of evolution and natural selection to economics. Its major feature is the emphasis on micro-foundations, that is a behavioural theory of firms based on the decision-making process described by Simon (1947 and 1955). Decision makers are characterised by ”bounded rationality” since they deal with incomplete information and uncertainty, two concepts which were originally discussed by Tintner (1941) and Alchian (1950). Bounded rationality implies both a capability gap in processing information and a competence gap in the ultimate representation of an agent’s environment, preferences, payoffs and problem-solving abilities. Evolutionary game theory has provided growing insights into the nature of rationality-bounded decision processes, characterised by adaptive, imitative and trial-and-error behaviours, and their implications for the selection mechanism at the aggregate level (Maynard Smith, 1982 and Samuelson, 2002).

At the micro level, bounded rationality implies the replacement of the neoclassical concepts of ”optimizing behaviour” and ”maximised profits” by the notions of ”satisfying” and ”positive profits” respectively. The uncertainty surrounding the economic environment leads firms to adopt “routines” and “rules of thumb”, namely decision rules applied routinely over a period of time, to search for better techniques of production (Nelson et al., 1976). An unsatisfactory level of current profits leads firms to revise their routines and engaging in the search process for a new technology, either as an internal R&D-driven activity or as an imitative process by technology transfer from other firms. Every time a new technology is found, firms perform a profitability check between the used

---

45 Rosenberg (1982) imputed the ”productivity” or ”Solow” paradox to a slow TC diffusion while Griliches (1994) pointed at measurement problems in the service sectors, that is where most of the productivity increase occurred. A critical view of the link between long waves and TC is provided by Mansfield (1983a).

46 Nelson (1995) and Witt (2003) describe the analytical building blocks of evolutionary theory.

47 ”Actual human rationality-striving can at best be an extremely crude and simplified approximation to the kind of global rationality that is implied, for example, by game-theoretic models” (Simon, 1959, p. 101).

48 Bounded rationality is also adopted as an explanation of industrial dynamics by a sociological research field called ecology of organizational populations (Hannan and Freeman, 1989; Carroll, 1997). This approach interprets industrial evolution as a process driven by entry, social legitimization and mortality of heterogeneous and inertial firms.

49 Routines are considered as ”as very close conceptual relatives of production "techniques" whereas orthodoxy sees these things as very different” (Nelson and Winter, 1982, p. 14).
technology and the potential one and choose consequently. Search is therefore a partial stochastic process characterised by locality and cumulativeness. The former indicates that the search for new technologies is likely to occur in the neighbourhood of the techniques already in use (Atkinson and Stiglitz, 1969). In turn, this implies the "cumulativeness" of TC, namely the relative importance of past experience of production and innovation upon future objectives. These two conditions make the concept of "path dependency" of TC (Section 5.1) as well as historical analyses of TC generation (Freeman, 1982) natural extensions of evolutionary literature. Moreover, evolutionary focus on a partial stochastic search process and dynamic industry effects (i.e. competition, growth and survival) represents a departure from the inducement mechanism described by the literature on "induced innovation", which stressed only the deterministic effect of a change in relative factor prices on inputs ratios in static terms. Nevertheless, market factors affect both the direction/intensity of a search process and the selection mechanism among different technologies, although in dynamic terms. As a result, "there is inherent plausibility in the Hicks inducement theory, biasing the long term direction of TC in a labour saving direction" (Freeman and Soete, 1987, p. 46). However, market conditions are not the only determinant of technology evolution since TC adjustment to a change in economic variables is constrained by its cumulative nature which "bind the scope for dynamic inter-factoral substitution" (Dosi, 1997, p. 1534). This implies evolutionary refusal of neoclassical production function in favour of the notion of technological paradigm (Dosi, 1982).

The concept of paradigm describes the framework upon which scientific knowledge relies on and develops over time (Kuhn, 1962). A technological paradigm refers therefore to the specific form of knowledge, the procedures and the basic system on which a particular activity is based. Moreover, it results from a complex selection process whose variables have a scientific, institutional and economic nature (Dosi and Grazzi, 2006). While the emergence of each technological paradigm represents a technological breakthrough, technological trajectories describe the rate and the direction of TC within each technological paradigm (Abernathy and Utterback, 1978), namely an incremental and gradual process of TC evolution which is enforced by a continuous process of learning and characterised by cumulativeness and irreversibility (Cantwell, 1999).

Technological paradigms are at the basis of evolutionary interpretation of the economic system whose main features are its dynamic/unstable nature and persistent heterogeneity. Their joint effect allows to extend the analysis to other sources of efficiency improvements besides the simple allocative one (Leibenstein, 1966). The focus on economic dynamics implies the rejection of mainstream methodological and functionalist conjectures related to the interpretation of equilibrium conditions as limit properties of some undetermined dynamics (Friedman, 1953). Studying the disequilibrium properties introduces a greater scope for the analysis of heterogeneous technologies and firms' behaviour, evolution, and survival probability. This introduces a fundamental aspect of evolutionary (or neo-Schumpeterian) theory, namely the idea that there is a continuous selection of firms by market mechanism over time. A crucial component of such representation is the occurrence of persistent heterogeneity among economic agents caused by bounded rationality and different learning processes (Lundvall and Johnson, 1994). Indeed, heterogeneity is strongly related to evolutionary...
The evolutionary search process provides also important insights into the analysis of technology diffusion (Metcalf, 1988). Indeed, heterogeneity and bounded rationality represent a strong departure from epidemic models of the 1960s since they allow the analysis of simultaneous available technologies and the effect of a competitive selection process on technology diffusion (Section 5.2).

This section has described the building blocks of the evolutionary theory of TC. Evolutionary contribution has established a more sophisticated representation of TC by stressing the relative autonomy of its internal dynamics from changes in the economic environment. Next section discusses this topic in more details. Moreover, evolutionary theory has provided several insights into the relationship between TC and economic growth. These contributions are discussed in Section 6.2.

### 4.2 A Two-way Explanation of Technological Change

Schmookler-related line of research on the inducement of TC was predominant in the 1960s in coherence with Keynesian national economic policies which emphasised more the effectiveness of "demand pull" policies rather than the importance of a "supply push" approach. In the 1970s, empirical research moved to the analysis of the simultaneous interaction of supply and demand factors. A dynamic two-way interpretation of the relationship between TC and economic factors replaced the "demand pull" version of the "induced innovation" theory, which was characterised by several shortcomings, such as an ambiguous approach to the Schumpeterian distinction between invention and innovation (Section 2.2). Indeed, rather than discussing the factors which determine a commercially successful innovations, "Schmookler analysed market demand forces as they influenced shifts in the allocation of resources to inventive activity" (Mowery and Rosenberg, 1979, p. 139).

Another point of criticism refers also to the imprecise concept of demand embraced by many "demand pull" studies. In particular, rather than discussing a systematic relationship between prices and quantities, they commonly adopt a very general definition of demand which included virtually all possible innovative determinants and, therefore, ruled out almost all other influences. Moreover, "demand pull" studies did not provide a solution to the identification problem which arises when it is necessary to distinguish between two parametric shifts, namely a movement of the demand curve and a movement along the demand curve. Indeed, only the former correctly supports the "demand pull" hypothesis while the latter may arise from a TC-induced downward shift in the supply curve which, in turn, generates an increase in demand. A similar identification problem refers to the distinction between those innovative factors which arise within the firm, such as output increase or a shift to a new technological base, from those which are mediated by the market (see, for instance, Myers and Marquis, 1969). "Demand pull" theories lacked also a satisfactory explanation of the inter-industry variations in technological performance as well as the timing and the direction of these innovations (Scherer, 1982; Walsh, 1984).

As a result of such analytical approach, there has been a small recognition of the way in which the growth of specialized knowledge has shaped and enlarged available technological capacities (Rosenberg, 1974b). Although economic forces affect the direction of TC, these effects take place

---

53 A recent line of research has applied this interaction to the study of "win-win" situations for the development of environmental standards (Porter and van der Linde, 1995 and Palmer et al., 1995).

54 Rothwell et al. (1974) and Spiller and Teubal (1977) provide evidence of the joint impact of demand-side and supply-side factors on the commercial success of innovation.
within the changing limits and constraints of a body of scientific knowledge growing at uneven rates among its component subdisciplines” (Rosenberg, 1976b, p. 270). Indeed, technology represents “a body of knowledge about certain classes of events and activities... it is not merely the application of knowledge brought from another sphere” (Rosenberg, 1982, p. 143).

Finally, at a very broad level, it is important to recognise that many important categories of human needs “have long gone either unsatisfied or very badly catered for in spite of a well established demand” (Rosenberg, 1976b, p. 267), which arises further doubts about market-driven analytical interpretation of TC evolution. The outcome of such attempts has been to leave TC inside a black box (Rosenberg, 1982). Rather than following this path, a more effective alternative seems to recognise the internal complementarities and cumulative patterns which emerge and characterise the technology sphere within the history of basic and applied science (Levin, 1977). However, this does not mean neglecting the effect of economic variables on the rate and direction of TC. Evolutionary approach has underlined that changes in relative prices and demand/supply conditions affect the rate and the direction of TC but only within given technological boundaries (Dosi, 1997). In general, it is correct to state that the ultimate incentives are economic, “but economic incentives to reduce cost always exist in business operations, and precisely because such incentives are so diffuse and general they do not explain very much in terms of the particular sequence and timing of innovative activity” (Rosenberg, 1969, p. 3). Modern industrial societies have been successful in applying systematised knowledge to the economic sphere (Kuznets, 1973, p. 249). Given these characteristics, a promising theoretical setting for the analysis of TC has to combine “the logic of scientific progress with a consideration of costs and rewards that flow from daily life and are linked to science through technology” (Rosenberg, 1982, p. 159), namely a model where “the relationship between technology and economics is really a "two way" interaction” (Vivarelli, 1995, p. 13).

This section has described the shift of economic research to the analysis of the simultaneous interaction of supply and demand factors for the description of TC evolution over time. These studies have mainly stressed the relative autonomy of technology internal dynamics. This issue has been discussed during the 1980s and extended through the concept of ”path dependence” (Section 5.1). Moreover, the relationship between TC and economic factors has been also addressed, although in a different perspective, by the debate on the correct TC-enhancing policy tools, and the relative analysis on appropriability and market structure (Section 5.3).

55Indeed, “...rather than simply referring to "lags" in the process, a useful theory of innovation must try to explain the varied length and distribution of such delays in the response to "needs"... such an analytic schema must explicitly consider institutional structures and dynamics, rather than static analyses” (Mowery and Rosenberg, 1979, p. 105).
5 Endogenising Technological Change (1980 - 1990)

The issue of economic growth gained its centrality again in the 1980s. The target in these years became the embodiment of recent analytical achievements into a more general theory of TC and growth. Directly related to the evolutionary theory and the dynamic representation of the relationship between TC and economic forces, "path dependence" theory stressed the cumulative path of technology evolution as the result of network externalities and increasing returns to adoption (Section 5.1). The 1980s witnessed also important contributions which extended technology diffusion models (Section 5.2) and related TC evolution to specific appropriability conditions and market structure (Section 5.3). Finally, the lack of empirical confirmation of (exogenous) growth models led to a new wave of growth studies which extended Solow’s framework (1957) by emphasising the role of knowledge and human capital as sources of long-run economic growth (Section 5.4).

5.1 Path Dependence

The "path dependence" hypothesis of TC (Arthur et al., 1983; David, 1987; Arthur, 1990) emerged during the 1980s as a result of the combination of different streams of economics literature. This line of research took the concept of increasing returns from the literature on "learning by doing" and stressed their role as a source of technological "lock-in". This concept refers to a situation where a market binds itself to a particular product, process, service or standard and high barriers to switching make the choice of a competing technology very difficult. "Path dependence" approach drew insights also into the "induced innovation" theory. In particular, the concept that localized induced TC may lower the elasticity of substitution in response to changes in market conditions (Wright, 1990) is seen as conducive to path dependent TC when the dominant force becomes the internal process of technology evolution (David, 1975, pp. 65-68). "Path dependence" theory elaborated the central concept of cumulativeness of TC evolution which derived from the evolutionary theory and the post-keynesian emphasis on the role of "positive feedback" and "cumulative causation". All these insights were then reinforced by several historical and industry studies (David and Bunn, 1988; David, 1993 and 1997; Stokes, 1994; Liebowitz and Margolis, 1995) and tested by means of simulation models of competing technology adoption. In particular, these models verify whether various orders of choice affect the final adoption shares under three different technological regimes, namely constant, increasing, and diminishing returns. The importance of early small events has been statistically represented in terms of time dependent probability outcomes (Polya, 1931; Arthur et al., 1987). Increasing returns in an industry provide therefore a self-reinforcement mechanism which could make the coexistence of incompatible technologies unstable and lead to the adoption of a single standard. However, increasing returns are a necessary, but not sufficient, condition for locking a system into a specific technology since other factors, such as adjustment costs, switching

---

56 An initial cost advantage may lead a technology towards a successful market outcome. A learning process and the occurrence of knowledge spillovers may, then, determine a new competitive advantage which deters the emergence of competing technologies (Katz and Shapiro, 1994).

57 History matters since future technological developments depend on the way the present state evolves (Foray, 1997). Models of regional growth which discuss the "success to breed success" tendency are provided by Kaldor (1970) and Dixon and Thirlwall (1975).

58 These models assume that an agent is perfectly acknowledged of each technology and its related payoff. However, he/she ignores which technology will be chosen by other agents.

59 The outcome indicated that "under constant and diminishing returns the evolution of the market reflects only a priori endowments, preferences, and transformation possibilities... under increasing returns, by contrast, many outcomes are possible. Insignificant circumstances become magnified by positive feedbacks to 'tip' the system into the actual outcome 'selected'. The small events in history become important..." (Arthur, 1989, p. 127).
costs between competing systems and the costs of maintaining parallel rival technologies, may lead to a market equilibrium with a technology characterised by a sub-optimal long-run potential. Positive network externalities refer to a situation where a given product is technologically more valuable to a consumer if other users adopt the same good or a compatible version of it (i.e. telephone and computer networks). Indeed, learning-by-doing and network externalities are commonly seen as those factors which locked users into the QWERTY keyboard system (the first six letters on the left of the topmost row of letters on the typewriter and now the computer keyboard), despite the alternative DVORAK keyboard layout resulted ergonomically more efficient. Three features characterise network externalities. Technical inter-relatedness stresses the need of users and producers for compatibility between adopted systems. Scale economies push towards a one-standard production through cost decreases. Finally, the quasi-irreversibility of investments underlines both the cumulativeness of the investment and the link with learning by doing. Moreover, network externalities can be direct or indirect, being the latter related to price reductions of related complementary products. Their occurrence introduces timing and strategic issues for both technology users and providers. On the demand side, network externalities introduce coordination problems even though technologies are supplied competitively. These problems can lead to excess inertia (users wait too long before adopting a new superior technology) or excess momentum (users rush to adopt an inferior technology). These two potential inefficiencies can be amplified in case of conflicting preferences or long information lags among users. Moreover, coordination becomes more difficult when there is a continuous entry of users in the market. On the supply side, network externalities affect how technologies are selected. This refers to the problem of product rivalry and a firm’s decision, cooperative or unilateral, whether making its technology compatible.

This section has discussed the determinants of technological path dependence by stressing the role of network externalities and standards. These factors exert also an influence on technology diffusion and market structure, two issues which are extensively assessed in the next two sections.

5.2 Technological Change and Diffusion: Some Extensions

Epidemic and probit models provided a theoretical framework for the analysis of TC diffusion by stressing the demand side of the technology market. Economic research broadened this analysis in several directions during the 1980s. Such extensions generally refer to all the three dimensions which affect technology diffusion, namely firm-level characteristics, the specific attributes of TC and industry-level variables.

Nevertheless, this result does not rule out the possibility that market forces and research effort would challenge the predominance of an inferior technology, even though this currently exploits increasing returns to its adoption and scale economies to its production.

On the effects of network externalities, see also Arthur (1994), Besen and Farrell (1994) and Liebowitz and Margolis (1994). Externalities can result in socially excessive adoption or “herding” effects when early adopters act randomly rather than on the basis of better information (Banerjee, 1992; Choi, 1997).

In this case, firms can use penetration prices and product preannouncements to install a base and favour the adoption of their technology. Standards, namely a specific technology chosen as the unique adopted in the market (Besen and Saloner, 1989; Postrel, 1990), communication and agreements between users, direct technology subsidies and public-funded R&D help to face coordination problems and to avoid technological lock-in.

Smaller firms have generally more incentive to enter the network through compatible products whereas stronger firms may prefer to enhance their market position by keeping incompatibility (Katz and Shapiro, 1986b).

Reviews of this literature can be found in Stoneman (1986); Feder et al. (1985) and Thirtle and Ruttan (2002).
At the firm level, diffusion studies drew a connection with the literature on investment by emphasising the complementarities between the generation and the adoption of technology (Cohen and Levinthal, 1989). On one side, extensions of the basic rank model provided a longer list of variables which affect a firm’s position in the rank, namely the level of gross returns a firm gains from technology adoption. Beyond R&D expenditure and firm size (Romeo, 1975), these variables are a firm’s ownership structure (Rose and Joskow, 1990), market share (Levin et al., 1987b), broader measures of adoption cost (beyond its purchase value) such as training or organizational costs (Stoneman, 1990), better geographic locations and skills availability (Ireland and Stoneman, 1985). On the other side, (new) epidemic models dealt with risk aversion, expectations, uncertainty and dynamic endogenous learning effects. In particular, different degrees of risk aversion among firms affect the value and timing of investment in new technologies and, consequently, the speed of diffusion (Geroski, 2000). Expectations may refer to the future price of a new technology, the cost of its acquisition (Balcer and Lippman, 1984) and the extent of demand, that is the number of its users (Oster, 1982). Uncertainty may be extended to both technological and economic variables (Just and Zilberman, 1983). Finally, a firm’s learning process over the actual performance of a new technology determines its adoption choice and, therefore, it affects technology diffusion (Lindner et al., 1979; Feder, 1980; Stoneman, 1980; Jensen, 1982).

A second extension to diffusion models considers a broader range of technology characteristics since improvements, modifications and adaptations to the requirements of various sub-markets increase the speed of diffusion (Rosenberg, 1972). The inclusion of the supply side into this setting allows to endogenise the continuous technological improvements originated by the suppliers. Empirical research assessed therefore the effect of vertical product differentiation, or technological improvements (Stoneman, 1989), horizontal product differentiation, or product variety (Stoneman, 1990), and network externalities on the diffusion process of new technologies (Stoneman, 1991; Midgley et al., 1992). Moreover, in line with the “induced innovation” literature, the adoption of a new technology is shown to be affected by the time path of its price and the price of those factor inputs which are intensively used by existing techniques (Hannan and McDowell, 1984). Finally, another distinction refers to the product or process nature of a new technology. Indeed, economic literature in recent years has focused on the diffusion patterns of process innovations since they allow an easier and more homogeneous representation of technology characteristics whereas empirical research on product diffusion has been mainly concentrated in the marketing literature (Mahajan et al., 1990). Within this group of models, only a small number of contributions have attempted to represent the selection and diffusion process of multiple technologies.

Finally, a third extension to diffusion models come from a deeper analysis of institutional variables, industry characteristics and strategic aspects involved in the adoption of a new technology. In particular, the typical sigmoid diffusion path of a new technology has been represented as the result of the interaction between supply-side and demand-side factors (Metcalfe, 1981; Stoneman and Ireland, 1983). Supply-side factors refer to several characteristics of the technology suppliers such as their price- and quantity-setting behaviour, their output capacity, the time structure of the supply.

---

65 Expectations can also affect diffusion through a shift of the optimal life of adopted technologies (Williamson, 1971). The role of expectations in the diffusion process is also discussed by Rosenberg (1976a) and Stoneman (1981).

66 A new wave of empirical research in technology diffusion and induced innovation appeared in the environmental literature where the search for and the adoption of a new technology has been related to government regulations (Johnson and Popp, 2003), fuel prices (Rose and Joskow, 1990), pollution abatement expenditures (Lanjouw and Mody, 1996 and Jaffe and Palmer, 1997) and energy prices (Newell et al., 1999; Popp, 2001 and 2002).

67 A diffusion model of product innovations has been proposed by Williams (1972).

68 The results obtained by these studies confirm the effect of the presence of multiple technologies on adoption decisions due to substitutions and complementarities in terms of costs and gains for the potential adopters (Ayres and Ezekoye, 1991; Stoneman and Kwon, 1994; Colombo and Mosconi, 1995).
costs and the market structure of the capital-good industry.\textsuperscript{69} An industry growth rate (Karshenas and Stoneman, 1993), government regulation and labour relations (Latreille, 1992) have been also shown to affect the diffusion of a new technology. On the demand side, evidence of the effect of market structure on technology diffusion has not provided any conclusive result. In particular, some studies find a positive relationship between market concentration and speed of diffusion (Davies, 1979) whereas others report a negative effect (Romeo, 1977) or no statistical effect at all (Karshenas and Stoneman, 1993). Strategic considerations in the determination of the timing of adoption, such as deterring imitation (or \textit{pre-emption}), emphasise a firm’s economic environment and entail stock and order effects as described by game-theoretic models (Dasgupta and Stiglitz, 1980a; Fudenberg and Tirole, 1985 and 1987).\textsuperscript{70} Strategic considerations arise also when diffusion is determined by technology transfer. An important aspect in this setting is the extent of appropriability conditions since “the relationship between property rights on an invention and the speed of its adoption is highly dependent on the kind of rent brought forth by the invention” (Tirole, 2000, p. 404).\textsuperscript{71} Indeed, market structure and appropriability conditions appear as crucial determinants of a firm’s innovative activity. Next section provides a review of the main relevant findings on this topic.

### 5.3 Appropriability and Market Structure

A substantial line of mainstream research has focused on appropriability conditions and market structure for explaining the rate and the direction of TC. The “induced innovation” approach has stressed the economic incentives which address a firm’s innovative investment (Section 3.1). A firm may choose between several types of knowledge inputs, being R&D commonly considered as the most important among them. This investment results in a knowledge output (Griliches, 1990) which, in turn, strongly affects productivity growth.\textsuperscript{72} Although R&D expenditure is a profit-motivated activity, some important features make R&D different from other types of investment. In particular, the skewness in the distribution of R&D outcomes, due to a mix of high variance of expected returns and a very low-probability associated to the highest payoffs (Scherer and Harhoff, 2000), influences a firm’s investment decision (Scherer \textit{et al.}, 2000) and makes the financing choice through capital market more difficult. Two further differences are related to the specific returns to R&D investment and its partial non-excludability. The former refers to the quasirents\textsuperscript{73} generated by R&D activity and it implies a potential “rent-stealing” effect (Mankiw and Whinston, 1986). Indeed, a firm which invests in R&D does not internalise the profit loss of its competitors on the product market and this may result in an over-investment in R&D (Lee and Wilde, 1980). In contrast, partial non-excludability arises the “appropriability problem”, which is likely to lead private firms investing too low in research activity (Dasgupta and Stiglitz, 1980b; Spence, 1984). There is some evidence that the appropriability effect prevails over the rent-stealing effect (Jones and Williams, 1984).

\textsuperscript{69}Moreover, the consideration of the supply side allows to assess the issue of technology diffusion policy through the definition of a welfare optimal diffusion path (Ireland and Stoneman, 1986).

\textsuperscript{70}These studies show a negative relationship between the number of technology users and the gross benefits from its adoption (Reinganum, 1981a and 1981b; Quirmbach, 1986). This results from the price increase of input factors and/or the price reduction of the final product (through a supply increase) due to technology diffusion.

\textsuperscript{71}Technology transfer between firms may be based on licensing (Gallini, 1984; Gallini and Winter, 1985; Katz and Shapiro, 1985b and 1986a; Kamien and Tauman, 1986; Rockett, 1990a and 1990b) or research joint venture (Ordover and Willig, 1985; Grossman and Shapiro, 1986b; Kamien \textit{et al.}, 1992).

\textsuperscript{72}Many studies have discussed the relationship between innovative investment and a firm’s productivity growth (Hausman \textit{et al.}, 1984; Griliches, 1986; Narin \textit{et al.}, 1987; Basberg, 1987; Crepon \textit{et al.}, 1998; Jorgenson and Stiroh, 2000; Loof and Hesmati, 2001; Toivanen \textit{et al.}, 2002; Hall \textit{et al.}, 2005; Parisi \textit{et al.}, 2006).

\textsuperscript{73}Quasirents represent the necessary incentives to perform sunk investments as R&D activity. They were originally described by Alfred Marshall (1920).
and, therefore, there is a private under-investment in the market compared to the desired social rates of R&D investment (Mansfield et al., 1977).

The appropriability problem paves the way to the role of government policy in promoting R&D (Goolsbee, 1998; Hall and Van Reenen, 2000; Klette et al., 2000; David et al., 2000; Martin and Scott, 2000) and the need of an appropriability-enhancing system which rewards and preserves a firm’s innovative investment (Jaffe, 1988). The relative importance of different tools for protecting innovation is related to the different degree of appropriability which, in turn, varies across time, markets, industries and technologies (Dosi, 1988).

Although other methods of supporting innovation exist, i.e. the contractual mechanism and the award system (Scherer, 1980, p. 458), the most widely used is the patent system (Jaffe, 2000). Nevertheless, patents generate a fundamental trade-off between the need of encouraging R&D and the emergence of a non-competitive environment which hampers knowledge diffusion and technology adoption (for a theory of optimal patent length, see Nordhaus, 1969). Such appropriability-enhancing system generates therefore a direct relationship between TC and monopolistic (product) market structure (Gilbert and Newbery, 1982).

The need of a departure from perfect competition to support innovation was originally proposed by Schumpeter (1942). Since then, economics literature has shifted from the view of market structure as an (exogenous) determinant of R&D activity to the recognition of a dynamic interaction between firm size, market structure and innovation (Kamien and Schwartz, 1975; Mansfield, 1962, 1981 and 1983b; Scherer, 1992). Such relationship is characterised by a different emphasis on either profit incentives or financial resources (Battaggion, 2000). On one side, (larger) firms in concentrated markets enjoy a higher degree of "appropriability" and, therefore, they are more able to exploit the returns to their innovations freed from competitive pressure (Cohen et al., 1987; Levin et al., 1985; Levin et al., 1987a). Second, they do not face financial constraints in their R&D investment since they rely on more financial liquidity due to both an easier access to external finance and larger internal funds to support costly innovative activities (Galbraith, 1952; Comanor, 1967). Third, large corporations in concentrated industries are characterised by a higher degree of diversification which helps them to deal with the uncertainty of innovation (Nelson, 1959).

On the other side, larger firms or firms in concentrated markets have scarcer incentives to innovate than firms in competitive markets since the latter may obtain a larger potential profit from their innovative activity (Fellner, 1951; Acs and Audretsch, 1990). Moreover, entrepreneurship studies have challenged Schumpeterian view by stressing the different technological and economic opportunities faced by big and small firms (Acs and Audretsch, 1987, 1988). In particular, a different management structure (Rothwell, 1989) and a less bureaucratic environment (Link and Bozeman, 1991) allow a higher responsiveness to innovative opportunities by small firms and new entrants in the industry.

---

74 A crucial determinant of this outcome is the role played by technological "spillovers" (Mansfield, 1985; Jaffe, 1986; Nadiri, 1993; Acs et al., 1994a). A large piece of literature has well documented their effects at the industry level (Bernstein and Nadiri, 1988 and 1989; Wolff and Nadiri, 1993; Sterlacchini, 1994), at the regional level (Jaffe et al., 1993; Audretsch and Feldman, 1996) and between different firms (Acs et al., 1994b; Audretsch and Vivarelli, 1996; D’Aspremont and Jacquemin, 1988; Los and Verspagen, 2000 and Cassiman and Veugelers, 2002).

75 Levin et al. (1984) considers as appropriability devices: patents, secrecy, lead effect, cost and time for duplication, learning curve, superior sale and service effort, economies of scale.

76 Schumpeter did not clarify whether the major determinant of innovative activity was either market share or firm size (Tirolo, 2000, p. 390). However, oligopolistic industries witness a larger average firm size and a higher probability that firms will undertake R&D investments (Chandler, 1977; Cohen and Klepper, 1996) than competitive industries, such as construction, which provide only a small contribution to overall R&D effort (Chen and Levin, 1989).

77 Scale economies and a different organization of work appear to support innovative activity of these firms (Mairesse and Mohnen, 2002). Indeed, innovative effort, measured by the employment of technical engineers and scientists, increases with market power (Scherer, 1965 and 1967a).

78 Several authors provide support to the entrepreneurship perspective (Baumol, 1968; Kenney, 1986; Evans and...
These alternative views reflect a different emphasis on the so-called efficiency and replacement effects (Arrow, 1962b). Both these effects refer to the pure incentive to innovation, namely apart from any cost or strategic considerations, of a firm which possesses a monopoly condition over the R&D activity. The replacement effect implies that the value of an innovation for the monopolist \( V^m \) is lower than the value of an innovation for the entrant \( V^c \), namely \( V^m < V^c \). The monopolist has a lower incentive to innovate than a potential competitor since the former replaces himself when the innovation takes place whereas the entrant gains a monopoly condition through its innovation. The efficiency effect introduces the threat of entry into the value assessment of an innovation. In particular, since a profit-maximising monopolist does not make less profit \( \Pi^m \) than two non-colluding duopolists \( \Pi^{d_1} + \Pi^{d_2} \) in the reference case of an industry characterised by a homogeneous product, the former, which may deter entry through an innovation, will gain a higher profit than a potential competitor entering the market. Therefore, \( \Pi^m \geq \Pi^{d_1} + \Pi^{d_2} \Rightarrow V^m \geq V^c \). The efficiency effect emphasises the (overall) industry profit reduction due to competition (Gilbert and Newbery, 1982).

A related line of decision-theoretic research has assumed the exogeneity of a firm’s environment (no strategic interaction) and discussed the optimality of R&D expenditures (Grossman and Shapiro, 1986a), search procedures (Weitzman, 1979) and funding criteria (Roberts and Weitzman, 1981).

A second stream of literature, based on a game-theoretic approach, has introduced strategic considerations into the analysis of a firm’s R&D investment and it has produced several models of patent-race. In these models, the race to the achievement of a specific innovative target (patents) allows to analyse the relationship between market structure, competition and innovative dynamics. A basic type of patent-race models does not consider a firm’s characteristics and focuses only on a firm’s research technology. This is characterised by a memoryless, or Poisson, process which implies that the probability of gaining the patent is a function of the current level of R&D investment rather than the stock of R&D accumulated over time (Loury, 1979; Reinganum, 1982 and 1984). This means that experience does not enter into the game and a firm’s R&D strategy is time-independent. The efficiency and the replacement effects, together with the nature of the innovation obtained, determine the Nash equilibrium of the game, namely the R&D investment \( x^* \) of each firm \( i \) such that \( x^*_i \) maximises \( V_i \) given \( x^*_j \). The more radical an innovation is, the more likely is the patent being gained by the entrant since the efficiency effect tends to zero \( (V^m = V^c) \) and there is not rent dissipation (Reinganum, 1983a and 1983b). In its extreme version, the continuous introduction of drastic innovations may result in a process of creative destruction characterised by a series of monopolies. In contrast, a cumulative innovation favours the persistence of monopoly (Gilbert and Newbery, 1984; Salant, 1984). In this case, a continuous stream of cumulative innovation may require a high level of R&D expenditure; in turn, this increases the probability of patenting for every unit of time and, thus, the efficiency effect overcomes the replacement one.

Several extensions to the basic patent-race models have been provided over time. One variation includes learning within the R&D process. This “experience” variable considers the effect of accumulated R&D stock, rather than its current investment level, on a firm’s patenting probability per unit of time. Experience appears able to narrow competition (Harris and Vickers, 1985) due to “e-preemption”, namely the first-mover advantage of the incumbent, which may gain the patent even though it has started the race only a short period before the competitor. However, a contrasting possibility is that the follower will “leapfrog” the leader by accumulating more experience and jumping ahead in the race (Fudenberg et al., 1983). This would justify the empirical evidence of competition at the R&D stage and may result from either information lags or a multi-stage patent-

Leighton, 1989; Geroski, 1989 and 1990; Jovanovic, 2001). Moreover, there is also evidence of the informal nature of innovative activity in small firms (Santarelli and Sterlacchini, 1990), which determines a downward bias estimate of their innovative propensity (Kleinknecht, 1987 and 1989 and Kleinknecht and Verspagen, 1989).
race setting. In the former, firms observe their rivals' R&D efforts with some delays. In the latter, jumps in experience determines a discontinuous variation of this variable.

Further extensions of the patent-race framework include the possibility of creating a research joint venture and sharing intermediate results through licensing (Grossman and Shapiro, 1987), allowing variable R&D intensities (Harris and Vickers, 1997) and heterogeneous technologies with different degrees of correlation (Holmstrom, 1982 and Bhattacharya and Mookherjee, 1986) and riskiness (Klette and de Meza, 1986).79

Finally, patent-race models have been adopted to represent the Schumpeterian process of creative destruction (Futia, 1980; Dasgupta and Stiglitz, 1981 and Reinganum, 1985) and, thus, they have represented an important reference for a series of growth studies during the 1990s (Section 6.1).

5.4 New Growth Theory in the 1980s

Since the mid 1980s, a new wave of growth studies originated from the lack of reconciliation between the framework depicted by Solow (1956 and 1957) and empirical evidence accumulated over time. Economic research followed new paths for explaining Solow’s residual. For instance, the relaxation of the hypothesis of perfect competition has allowed an assessment of the relationship between economic growth and market structure (Hall, 1988) whereas an emerging literature on real business cycle has increasingly related economic cycles, uncertainties and growth (Domowitz et al., 1988; Benhabib and Farmer, 1994). However, the need to explain both the non-decreasing trend in per-capita growth and the absence of income convergence between countries pointed directly at the insatisfactory nature of exogenously-driven explanations of long-run productivity growth (Maddison, 1979; Baumol, 1986). New growth theory has focused on the effect of knowledge and TC on the ability of an economy to overcome the problem of diminishing returns and, thus, maintaining a positive per capita growth in the long-run. The aim of converting the level of technology into an endogenous element was actually anticipated by Arrow (1962a) model of "learning by doing" which turned out to be a prototypical "AK model" of economic growth (Rebelo, 1991).80

"Linear endogenous" growth models adopt a broad concept of capital which includes both physical and human factors (Knight, 1935 and 1944). These models implicitly assume therefore that all factors of production can be accumulated. In this setting, the inclusion of non-reproducible factors, such as raw labour, modifies the property of constant returns to scale and it results in a model characterised by increasing returns. In turn, this constitutes a major obstacle to a general equilibrium framework since increasing returns imply unconstrained production profits. Growth theorists solved this problem by the adoption of spillovers of the Marshallian type (Marshall, 1920; Stigler, 1939). Indeed, spillovers provide the basis for a production function with constant returns to scale at the individual level, namely for a given stock of knowledge, and increasing returns to scale at the aggregate level which, in turn, determine an overall improvement in productive capabilities (Romer, 1987). Such specification allows therefore the balance between the need of maintaining the assumption of perfect competition (for the sake of modelling simplicity) and representing a situation in which a positive per-capita growth occurs.

Since externalities and knowledge spillovers arise from activities such as learning by doing, R&D and education, new growth theory alternatively considered variables as knowledge (Romer, 1987).

79 On the issue of risk, see also Lambert (1986) and Holmstrom and Costa (1986).
80 The "AK model" is described by the following aggregate Cobb-Douglas production function \( Y = AK \). The similarity between the two models is verified in the specific case of "learning elasticity" equal to 1 (Solow, 1997), which makes Arrow's (1962a) model linear with respect to capital. Despite its weakness, this setting served as a basic reference for more sophisticated linear models of factor accumulation (Benveniste, 1976; Eaton, 1981; Jones and Manuelli, 1990) where the absence of decreasing returns to capital allows a positive growth rate in the long-run.
1986) and human capital (Lucas, 1988) as further input factors into the production structure\textsuperscript{81}. Although labelled as models of endogenous TC, these contributions aimed at endogenising the effects of TC, rather than TC itself, since the lack of micro-foundations did not allow to depict the process of knowledge production within the framework of these models. Although Romer (1986) and Lucas (1988) shared the same basic framework, namely a two-sector model augmented by a knowledge-related variable, they followed different modelling strategies. Beyond the production sector, Romer (1986) assumed a research sector which produces knowledge by physical investments while Lucas (1988) described the generation of knowledge through accumulation of human capital in the education sector\textsuperscript{82}. A two-sector framework allows to depart from Arrow (1962a) by placing the knowledge creation away from physical investments in the production sector. In Romer (1986), investments in the research sector are characterised by diminishing returns whereas the production sector exhibits increasing returns due to externalities\textsuperscript{83}. In Lucas (1988), there are non-decreasing returns to human capital investment since this is accumulated during leisure time which, in turn, is assumed to be constant\textsuperscript{84}. The human capital accumulation equation and the production function in Lucas’ (1988) two-sector economy run as follows: $h^o = \phi h(1-u) \implies Y = AK^\beta (uhL)^{1-\beta}$ where $u$ is the proportion of total labor spent working, $\phi$ is the productivity parameter of the human capital investment and $h$ represents the stock of “human capital”\textsuperscript{85}. The linearity of the human capital equation and the absence of restrictions on it make unbounded growth possible. In turn, they cause the benefits arising from human capital, amplified by Marshallian externalities, to represent the driving force of long-run economic growth.

Beyond theoretical achievements, a purely empirical growth-accounting approach emerged during the 1980s as a result of a wider availability of empirical data (Nelson, 1981; Summers et al., 1984). This approach related the origin of productivity increase to education, R&D and other technology variables (Mansfield, 1972)\textsuperscript{86}. However, these empirical analyses have suffered from both shortcomings in the measurement of the rate of quality improvement and the perfect competition assumption. The former is especially relevant in the case of ICTs (information and communication technologies) and it creates a bias in the estimation of TC contribution (Maddison, 1987). The latter implies that factors are paid their marginal products, a misleading assumption in the case of oligopolistic market structures.

Many of these challenges, common to both theoretical and empirical analyses in the 1980s, were tackled in the following years. However, while the 1970s were considered as a lost decade for mainstream research into the determinants of long-run economic growth, the 1980s set its agenda towards a deeper understanding of the relationship between growth and TC.

\textsuperscript{81}The difference between knowledge and human capital is that “knowledge refers to society’s understanding about how the world works. Human capital refers to the resources expended transmitting this understanding to the labor force.” (Mankiw et al., 1995, p. 298). Although the importance of these two concepts was already recognised in previous economics literature (Schultz, 1961; Krueger, 1968; Bardhan, 1993), there were not formal attempts to relate knowledge and human capital to economic growth until the 1980s.

\textsuperscript{82}These two models gained many insights from Uzawa’s (1961b, 1963 and 1965) attempts to endogenise TC in a two-sector growth model. A comment to such setting is provided by Inada (1963).

\textsuperscript{83}Non-decreasing returns to scale require a departure from the hypothesis of perfect competition, an issue which will be explicitly addressed in the successive decade (Section 6.1), and, thus, do not allow the standard factor-share measure of the contribution of capital and labour in a production function framework.

\textsuperscript{84}The hypothesis of constant returns to scale is crucial for this model. Indeed, increasing returns to scale would make the growth of capital infinite in a finite time whereas decreasing returns to scale would make this model unable of generating permanent growth (Mulligan and Sala-i-Martin, 1993; Solow, 1994).

\textsuperscript{85}Human capital accumulation affects the capital-labour ratio through changes in productivity levels and complementarities between skills and capital (Griliches, 1969).

\textsuperscript{86}The “knowledge production function” (Griliches 1979; Griliches and Lichtenberg, 1984), namely a production function augmented by a knowledge stock measure, has provided direct evidence of the relationship between GDP and R&D investment (Griliches, 1986; Griliches, 1995).
6 Recent Approaches to Technological Change (1990 - )

Historical events, such as the emergence of new ICTs, a faster trend towards global economic integration, and a policy shift aimed at the building of a modern “knowledge society” (European Union, 2005), have deeply reinforced the recognition of the crucial role of TC in promoting economic growth well beyond the boundaries of economics departments.

This section focuses on recent analyses of the relationship between TC and growth from the two main competing approaches in economics literature, namely neoclassical endogenous growth theory (Section 6.1) and evolutionary analysis of TC (Section 6.2). As discussed in Section 4.1, the emergence of the latter has reflected the dissatisfaction toward key aspects of mainstream approach, such as the representation of TC as a "black box”, namely the lack of micro-foundation in the analysis of the sources of TC, and economic growth as a ordered, steady state concept.

Indeed, apart from “ideological-like” arguments, evolutionary focus on complex dynamics, uncertainty and heterogeneity has led to a more variegated representation of actual economic process at a cost of a weaker theoretical robustness. On the contrary, neoclassical tradition has been a productive source of empirical research at the cost of sacrificing some degree of adherence to the real working mechanism of the economic system.

6.1 New Growth Theory in the 1990s

Growth studies in the 1980s stressed the role of knowledge and human capital in promoting economic growth. At the beginning of the 1990s, the challenge consisted of providing a direct explanation of long-run growth through a model where TC itself is endogenous, namely a model which stresses both the role and the incentives of innovative investment since "a story of growth that neglects technological progress is both ahistorical and implausible" (Grossman and Helpman, 1994, p. 26).

Beyond increasing returns to scale (Backus et al., 1992; Burnside, 1996), new growth theory started to consider some additional factors, such as imperfect competition, international interdependence and incomplete appropriability, to provide a better explanation of the determinants of knowledge investment. In doing so, this line of research drew on and combined different streams of literature to provide original models of endogenous TC-based growth. These include the research into the determinants of the rate and the direction of TC, learning by doing and the debate on knowledge externalities, appropriability, uncertainty and market structure during the 1980s.

Growth models in the 1980s framed knowledge investment within a standard perfect competitive economy. In fact, this was at odds with a well-established literature on the determinants of R&D which, on the contrary, stressed the incentive mechanisms behind this investment (Dasgupta and Stiglitz, 1988). New endogenous growth theory moved therefore towards an economy featured by monopolistic market structures where innovative investments represent an endogenous equilibrium response to Schumpeterian profit incentives.

Evolutionary approach reflects the Hegelian view, namely a dialectic process between chance and necessity. On the contrary, mainstream growth theory refers more to a Newtonian vision, namely a view of the world as a clockwork in which enough information about the present state allows a perfect prediction of its future states (Verspagen, 2005). The following authors provide an overview of the contributions to new growth theory: Lucas (1993); Pack, (1994); Temple (1999); Solow (1999) and Verspagen (1992).


Young (1993a) and Jovanovic and Nyarko (1996) combine learning by doing and growth in their model. The former stresses the interdependence between research activity and production experience. The latter discuss the implications of learning by doing for productivity increase and technology choice.

This shift challenged the neoclassical assumption that policy can “affect the level of economic activity but not the
Romer (1990) discusses an economy composed by three sectors. The research sector uses human capital and the existing stock of knowledge to produce new knowledge, which is sold in the form of designs for new producer durables to the intermediate-good sector. This sector uses the designs and foregone output to produce an increasing variety of capital goods. Finally, these goods are used in the final-good sector which is characterised by the following Cobb-Douglas production function:

\[
Y(H_Y, L, x) = H_Y^\alpha L^\beta \int_0^\infty x_i^{1-\alpha-\beta} di
\]

In this sector, production inputs are raw labor \(L\), the share of human capital devoted to the final-good sector \(H_Y\), and an indexed classification of capital reflecting the (infinite) variety of capital goods available \(x(i)\). This production function, which is taken from Ethier (1982) and based on the "love of variety" idea of Dixit and Stiglitz (1977), exhibits constant returns to scale and it depends positively on the number of different capital goods used to produce output. Economic growth is therefore a function of a larger variety of capital goods. The monopolistic structure in the research sector creates the market incentives for undertaking innovative investment. Knowledge is a partially excludable and non-rival good. The former reflects the feature of the physical outcome of R&D activity, namely the designs for new producer durables, which is protected by patent rights and it responds to market incentives. The latter indicates that R&D generates spillovers since a new stock of freely available knowledge contributes to expand the knowledge frontier in the research sector and to increase human capital productivity. Therefore, constant returns to R&D are possible, the incentive to innovate remains positive and knowledge grows forever.

While this model describes growth as a process of horizontal differentiation, Grossman and Helpman (1991d) and Aghion and Howitt (1992) capture the Schumpeterian ideas in a framework where repeated product improvements (vertical product differentiation) determine a firm’s monopoly power in supplying the best-practice capital good to the final good sector. Unlike Romer’s (1990) setting, monopoly power is only temporary since each innovation supersedes the previous one (quality ladder) and new innovations are built on the previous ones. Although horizontal and vertical innovations provide the same insights into the determinants of long-run growth, different normative implications arise (Grossman and Helpman, 1991d). First, the former imply a constant R&D underinvestment whereas this is not true for the latter. Indeed, vertical differentiation describes a setting very similar to patent-race analyses in terms of both a firm’s ability to estimate the probability of gaining the patent, conditional to its level of R&D spending, and the coexistence of appropriability problems and business-stealing effects (Aghion and Howitt, 1996). Moreover, a model featured by repeated product improvements allows to verify the importance of the Schumpeterian process of creative destruction. In turn, market imperfections and creative destruction emphasise the role of a profit-seeking entrepreneur in generating economic growth by her ability to deal with continuous technological and economic opportunities (Schultz, 1975; Shane, 2001).
While models of vertical differentiation depicts TC as a continuous cumulative process of quality improvements, a different stream of research has underlined the contribution of general purpose technologies (GPT) to economic growth (Bresnahan and Trajtenberg, 1995). These are basic technologies (radical innovation) which find use in many different sectors within the economy. GPT contribute to economic growth by increasing the returns to those R&D activities which incorporate them into distinct production sectors (Helpman, 1998)\textsuperscript{98}.

However, neither creative destruction nor GPT-based models allow coexistence and complementarities among different technologies (Young, 1993b). The recognition of such heterogeneity, together with the importance attached to international technological diffusion (Keller, 2004), represent the key features of TC-based models stressing the role of trade in promoting growth at the country level. Rather than identifying a direction of causality between trade and TC (Posner, 1961), these models underline the interaction between these two processes. A series of contributions by Grossman and Helpman (1990a; 1990b; 1991c; 1991e) have provided a useful synthesis between TC and trade literature. By dropping some unrealistic assumptions, such as the existence of a common production function across countries, they advocate the growth-enhancing effect of trade by stressing more the resulting technological upgrading rather than the (mainstream) resource-reallocation argument. The former is the result of an easier transmission of technological information, greater economies of scale and a greater amount of competition. Moreover, empirical evidence tend to weaken the latter since the uneffectiveness of international capital flows in equalising income across countries. In fact, this calls for a persistent heterogeneity in human capital accumulation (Lucas, 1990; Stokey, 1991), technological investments (e.g. discovery vs. imitation: Grossman and Helpman, 1991b) and Schumpeterian dynamics (Segerstrom et al., 1990).

The evolution of the literature on growth theory and a wider availability of empirical data have stimulated a new wave of empirical research on the determinants of intertemporal cross-country growth variability. However, at least until the first half of the 1990s, most empirical work tested the convergence hypothesis within the Solow’s (1957) framework rather than assessing the contribution of new endogenous growth theory. This has led research to focus more on the aggregate returns to capital (Barro and Sala-i-Martin, 1992; Sala-I-Martin, 1997) rather than on the strength of capital externalities (Section 5.4). Moreover, there has been a preference towards a mere statistical-driven approach rather than a theoretically-sound analysis. This has resulted in a series of empirical results, mostly based on aggregate cross-country studies, fragile to a sensitive analysis. For instance, Mankiw et al. (1992) estimated the following production function:

\[ Y_t = K_t^a H_t^b A_t L_t^{1-a-b} \quad \implies \quad Y = K^{1/3} H^{1/3} L^{1/3} \quad (2) \]

where \( Y \) is output, \( K \) physical capital, \( H \) human capital, \( A \) a constant technology parameter and \( L \) raw labour. Their cross-country analysis provides an empirical outcome, proxied by a production function where the power of each input is 1/3, which leads to the conclusion that "there are not substantial externalities to the accumulation of physical capital" (p. 432). They neglected therefore the validity of the Romer-Lucas contribution and advocated the ability of an augmented Solow growth model to predict conditional convergence\textsuperscript{99}. However, a crucial assumption of this setting

\textsuperscript{98}In general, GPT-based models of economic growth draw on the concepts of technological paradigm from the evolutionary approach to TC (Dosi, 1982) and both long waves (Jovanovic and Rob, 1990) and cyclical growth (Jovanovic and MacDonald, 1994) from Schumpeterian tradition.

\textsuperscript{99}Similar support to an augmented version of the neoclassical growth model is provided by Hayami and Ruttan (1970), Kawagoe et al. (1985) and Mankiw (1995).
is that countries share the same technological level. In fact, this is at odds with the persistence of heterogeneous technological capabilities among countries over time.

Many empirical studies have provided evidence of the relationship between economic growth and several economic variables\textsuperscript{100}. In general, three results appear robust in the empirical literature (Levine and Renelt, 1992). First, a positive and robust correlation between the average share of investment in GDP and growth rates (Wolff, 1991; De Long and Summers, 1991 and 1993; De Long, 1992; De Long \textit{et al.}, 1992; Helliwell and Chung, 1992). Second, a positive support to the conditional convergence hypothesis, namely a negative relationship between the initial level of per capita income and subsequent economic growth when human capital accumulation is included in the analysis (Barro, 1991). Third, a positive correlation between the investment share and the ratio of trade to output (Grossman and Helpman, 1990b; Rivera-Batiz and Romer, 1991).

A recent wave of empirical research has adopted more sophisticated econometric techniques and it has attempted to verify the theoretical insights of endogenous growth theory by focusing on the channels of technology diffusion. This has led to a series of contributions which have underlined the role of increasing returns in explaining agglomeration economies (Ciccone and Matsuyama, 1999; Krugman, 1991b) and their effects on productivity growth (Ciccone and Hall, 1996; Ciccone, 2002). However, most empirical work has dealt with the international transmission of R&D spillovers\textsuperscript{101}. The importance of technological spillovers has pushed research towards the identification of the actual channels of technology transmission, such as technological cooperation and employee mobility (Griliches 1992) and the international trade of capital-embodied TC (Eaton and Kortum, 2001)\textsuperscript{102}.

Moreover, the distinction between the effects of domestic and foreign innovative activity, induced by knowledge spillovers, has led to the assessment of the complementary role of technology creation and diffusion as drivers of a country’s growth (Eaton and Kortum, 1999). This concept refers to both a country’s absorptive capacity (Cohen and Levinthal, 1989), namely the ability to adopt/complement foreign knowledge and, thus, catch-up countries at the technological frontier (Fagerberg, 1995), and the role of institutions as preconditions of economic growth (Ruttan and Hayami, 1984)\textsuperscript{103}.

The recognition of the importance of institutional change and structural change is mainly rooted in development economics and technological-gap theory, which have discussed the conducive role of a "national innovation system" to a sustained economic growth (Abramovitz, 1986)\textsuperscript{104}.

Exactly these issues have witnessed a convergence of research interests over time between mainstream economics and evolutionary theory of TC and growth, whose main achievements are discussed in the next section.

\textsuperscript{100}Economic growth is found dependent on public spending (Barro, 1990), population growth (Nerlove, 1974; Griliches, 1974; Becker \textit{et al.}, 1990; Kremer, 1993), tax rate (King and Rebelo, 1990), public capital (Shioji, 2001), income distribution (Kuznets, 1955; Galor and Tsiddon, 1996), poverty (Sen, 1983), natural resource constraints (Grossman and Krueger, 1995) and savings rate in either a Ramsey (1928) - Cass (1965) framework or an Overlapping Generations (Diamond, 1965) setting (Abel and Blanchard, 1983).

\textsuperscript{101}These analyses adopt a production function framework which allows to disentangle the returns to the R&D stock of the unit of observation (firm, country or sector) and to the other units (Lichtenberg, 1992; Brecher \textit{et al.}, 1996; Coe \textit{et al.}, 1997; Zigic, 1998; Bin and Jianmao, 1999; Fritsch and Franke, 2004; Lumenga-Neso \textit{et al.}, 2005).

\textsuperscript{102}Empirical research has dealt with the measurement of R&D spillovers through different weighting schemes based on a country’s trade flows (Coe and Helpman 1995), intersectoral flow matrices (Verspagen, 1997), foreign direct investments (Lichtenberg and Van Pottelsbergh, 1998) and monte-carlo simulations (Keller, 1998).

\textsuperscript{103}Indeed, a country’s long-run growth appears to converge to its own steady-state if the analysis controls for fixed-country effects, such as the role of institutions (Islam, 1995).

\textsuperscript{104}The former refers to property rights (Clague \textit{et al.}, 1996) and contract choice (Stiglitz, 1974; Otsuka \textit{et al.}, 1992). The latter indicates the historical process of convergence among countries (Broadberry, 1994 and Verspagen, 1995) and the sectoral shift from agricultural and manufacturing within countries (Jorgenson, 1961; Ranis and Fei, 1961; Echevarria, 1995 and 1997).
6.2 The Evolutionary Approach to Technological Change and Growth in Recent Years

The foundation of an evolutionary theory of TC during the 1970s paved the way to several lines of research on the determinants of industry evolution and economic growth. The evolutionary approach has interpreted the persistent heterogeneity in these phenomena as the collective appearance of some emergent properties. These properties refer to the stochastic nature of the process of industry entry and to the dynamic attribute of the selection mechanism among firms and technologies. A continuous expanding set of technological opportunities defines industry evolution while different technologies determine a heterogeneous profitability among firms and, ultimately, different investment capacity, survival probability and growth. In turn, these aggregate facts result from non-equilibrium interactions among bounded-rational and heterogeneous economic agents (Metcalfe, 1995).

The coevolution between micro and macro-economic variables (Nelson, 1995; Malerba and Orsenigo, 1996b), together with the rejection of the production function neoclassical framework, has pushed most evolutionary research towards the adoption of simulation models of TC and growth. This methodology consists in the empirical identification and simulation of some key stylized facts by the model. In particular, the variables of interest, such as the output value, the factor-price ratio or the rate of capital accumulation of an economy, are endogenised through calibration procedures.

The main distinction among evolutionary models is related to the assumptions about the behavioural properties of the economic agents and the environment in which their interactions occur. The former refer to the degree of forward-looking rationality of firms. The latter distinguish the different hypotheses behind the persistence of stochastic entry into the market (i.e. “open-ended” dynamics) and the identification of the agents which are actually able to innovate.

Several empirical works on industrial organization have provided the “stylised facts” described by evolutionary models.

At the firm level, heterogeneity is due to both accountable market characteristics and learning variables. The former refer to pricing behaviours, investment strategies, profitability and market shares (Jensen and McGuckin, 1997). The latter describe widespread and persistent variability in routines (Garud and Rappa, 1994), technological capabilities (Freeman, 1994; Geroski et al., 1997) and learning (Winter et al., 2000).

At the industry level, entry and exit patterns are found persistent over time (Dunne et al., 1988; Klepper, 1996). Since the survival rate of most entrants is low, entry rates are higher than market penetration rates and, thus, a skewed firm-size distribution appears stable over time (Geroski, 1995).

---

105 Evolutionary game theory has discussed non-equilibrium interactions (Levinthal, 1997) in both a deterministic and stochastic settings (Friedman, 1991 and Weibull, 1995). This research stream focuses on a micro-founded interpretation of some aggregate concepts, such as a demand curve or GDP growth (Silverberg et al., 1988). A series of models on "artificial worlds" (Lane, 1993a and 1993b) has dealt with the relationship between the out-of-equilibrium aggregation of individual conditions and the emergence of aggregate statistical regularities (or “metastability”). Silverberg and Verspagen (1994) provide a long-run model of industrial dynamics which adopts this framework.

106 On one side, Winter et al. (2000 and 2003) provide an analytical framework which does not assume a firm’s behavioural rationality. On the other side, equilibrium conditions and a higher level of agents’ rationality characterise the model of Jovanovic (1982). Indeed, survival probabilities and growth rates are determined by an equilibrium path of optimal adjustments dependent on a firm’s revealed performance. Such path has been also modelled through both a Markov process (Hoppenhain, 1992) and Markov-perfect Nash equilibria (Ericson and Pakes, 1995).

107 For instance, Bottazzi et al. (2001) provide a model where both entrants and incumbents continuously introduce innovation whereas Dosi et al. (1995a) limit such possibility only to entrants. Extensions of the basic evolutionary setting have also targeted the relationship between market structure and the competitive process (Utterback and Suarez, 1993) and the distinction between imitation and innovation (Andersen, 1994).

108 Indeed, the distinction between tacit and codified knowledge (Malerba and Orsenigo, 2000b) makes the existence and evolution of various types of learning process an important determinant of heterogeneity among firms (Malerba, 1992; Dosi et al., 1999).
Indeed, industry-specific characteristics, such as the extent of scale economies, capital intensity, and the degree of innovativeness of the industry affect a firm’s technological capability as well as survival rates (Audretsch, 1991, 1995a, 1995b; Doms et al., 1995). Evolutionary-based industry studies have provided growing insights into the historical evolution of particular industries (Malerba et al., 1999) by disentangling the effect of exogenous variables and chance events on an industry’s early evolution and maturity (Klepper and Graddy, 1990). In particular, these studies have represented an industry life cycle through different "transitory" phases. In turn, this enables an assessment of the role of endogenous shakeouts (Klepper, 1997; Ufuah and Utterback, 1997) and spinoffs (Hirschman, 1970; Klepper, 2005) for innovative dynamics and industry evolution.

Economics literature has discussed technological differentials across sectors by providing different taxonomies based on some specific industry characteristics. In particular, from the Schumpeterian distinction between an "entrepreneurial regime" and a "routinized regime" (Winter, 1984), the analytical consideration of the role of technology flows has led to the widely-used quadripartite Pavitt (1984) taxonomy up to the broader concept of "sectoral systems of innovation and production" (Malerba, 2002). Moreover, an important explanation of heterogeneous innovative rates across sectors is provided by the concepts of "technological regime" and "market regime". The former allows to explain the observable variety in industry structure and technology evolution across industries (Malerba and Orsenigo, 1996a; Audretsch, 1997). The latter refers to the process of market interaction among firms (Freel, 2003).

At the country level, evolutionary models describe economic growth as a TC-driven phenomenon. This is characterised by a complex alternate process of convergence and divergence over time (Silverberg and Verspagen, 1995) and persistent technological heterogeneity across different economies (Chiaromonte and Dosi, 1993). Such heterogeneity gives room to a country’s choice between innovative versus imitative strategies which, in turn, lead to the possibility of "catching up" or "falling behind" (Verspagen, 1991). A non-linear growth process results from the continuous emergence of technological paradigms (Section 4.1), which allows to avoid aggregate decreasing returns to scale in the long-run. Each paradigm affects a country’s economic growth through the co-evolution of its constituent factors. Evolutionary models have emphasised therefore the growth effect of the stochastic but clustered arrival of new technologies (Silverberg and Lehnert, 1993), the speed of its diffusion (Conlisk, 1989), the role of institutions and government policy (Dosi et al., 1990; Geels, 2004), the nature of interactions among bounded-rational agents (Fagiolo and Dosi, 2003) and the role of demand.

---


110 Dosi et al. (1988); Day and Eliasson (1986); Silverberg and Soete (1993) and Fagerberg (1994) provide a review of the evolutionary literature on economic growth.

111 The importance of demand is a typical feature of post-keynesian studies which have investigated the relationship between TC and growth (Scott, 1989; 1992a and 1992b; Pasinetti, 1993).
7 An Agenda for Future Research

This survey has discussed the evolution of the literature on TC over time. Since its origin, this line of research has aimed at a deeper understanding of the ultimate reasons of positive long-run economic growth. The structure adopted in this paper has allowed to map the evolution of this literature and to stress the intertemporal linkages between different areas of research.

Several complementarities and bridging factors characterise the literature on TC. The former refer, for instance, to the analysis of the determinants of productivity growth. This has followed complementary paths during the 1960s in accounting the different quality of production inputs, namely the streams of literature on embodied TC (Section 3.2) and learning (Section 3.3). Moreover, mainstream and evolutionary literature deal with economic growth nowadays by emphasising complementary approaches (macro vs. micro), analytical tools (i.e. intertemporal maximisation and production function vs. simulation studies) and research interests (equilibrium vs. disequilibrium). It is possible to find also several examples of bridging factors between different streams of research over time. For instance, the original learning by doing setting of Arrow (1962a) has led to the advancements of new growth theory during the 1980s such as the human capital model of Lucas (1988). A further example refers to the evolution from mono-casual explanations of TC during the 1960s (Section 3.1) to a two-way interactive theory of the relationship between TC and economic factors during the 1970s (Section 4.2).

As pointed out by Abramovitz (1993, p. 217), “the seeming major contribution of tangible capital accumulation to nineteenth-century growth was the consequence of scale-dependent and capital-using TC. The large twentieth-century contributions of education and R&D conceal technology’s new intangible capital-using bias”.

Shifts in the determinants of long-run economic growth have resulted also in a noticeable change of the research agenda. At the micro level, there is an increasing attention nowadays to the ultimate reasons of productivity growth through a direct analysis of the optimal way to encourage TC and a micro-founded explanation of the dynamics internal to the "black box". This refers, for instance, to theoretical and empirical contributions on the role of policy and institutions in supporting TC and the effect of different economic and technological variables on measures of innovative activity at the firm level. At the macro level, the growing process of economic globalization and sector specialization calls for a greater role of knowledge spillovers, economies of scale and international trade in explaining productivity growth and technological catch-up. Increasing integration between countries increases the speed of technology adoption which, in turn, affect social and economic variables. The pervasiveness of this process poses a challenge to traditional national-based interpretation of economic phenomenons and it strongly requires the enlargement of traditional theoretical frameworks by the assessment of TC dynamics, also in economics fields which do not have traditionally dealt with the effects of technology and innovative activities112. Beyond economic integration, the increasing service specialisation of developed economies and the role of venture capital as a financing source of innovative ideas provide new opportunities for entrepreneurship and economic growth (Gompers and Lerner, 2001a and 2001b; Audretsch et al., 2006). It is necessary therefore to deal with the specific patterns of TC in these sectors since they will increasingly affect a country’s overall economic growth.

The main objective appears therefore to provide original insights into a new research agenda by clarifying the underlying mechanisms which relate TC and growth. This approach has to be consistent with the occurrence of a persistent sector and regional heterogeneity in the innovative

\[112\] An example refers to the trade-off between unemployment and inflation described by the Phillips curve. This relationship has flattened considerably in the last 15 years as a result of a greater economic integration among countries (Rogoff, 2006).
process and the simultaneous presence of different kinds of TC at any time.

A necessary analytical step appears the integration of different lines of research whose complementarities may prove essential in building a unified coherent picture of such relationship. Therefore, an agenda for future research would focus on two major points.

The first challenge is to integrate the results obtained so far by the literature on endogenous growth theory, the determinants of TC, evolutionary economics and path dependence. Although the disagreement on the essential nature of the economic system and its actors can not be reasonably overtaken, neoclassical growth theory has incorporated some central issues of the evolutionary approach over time, such as the departure from the perfect competition hypothesis, the representation of R&D and technology as stochastic phenomena and the role of knowledge flows for long-run growth. It seems possible therefore to move towards further generalizations of the neoclassical setting which address directly the issues of heterogeneity and bounded rationality. The former may be discussed through the identification of different agents in the economic system rather than the mere analysis of a representative one (i.e. a firm). Such analysis may result from the empirical identification of a firm’s conflicting preferences, strategies, and some relevant specific features applied to innovative activities. A first step for addressing the issue of an agent’s rationality can be a move towards its stochastic representation, similarly to the representation of R&D and technology already introduced in the literature.

In general, it appears plausible to develop a macro-economic model where both the rate and the direction of TC are endogenous. This can be done by integrating market factors and demand-pull variables on one side, and technology-push and path-dependence insights on the other side (Landes, 1994). The assessment of the relative effect of such complementary forces may gain from some evolutionary insights as the emphasis on history, institutions and technological heterogeneity in terms of different technological paradigms (Dosi, 1982), sectoral systems of innovation (Malerba, 2002) and technological regimes (Breschi et al., 2000). The emphasis on micro-foundations has led evolutionary research to target a more effective representation of TC in terms of three complementary levels of analysis of technologies (Winter, 2006). Research on these complementary levels is still at the beginning. There is therefore room for further theoretical and empirical enlargements. The first level refers to the definition of some bodies of problem-solving knowledge, and it involves the nature of knowledge upon which TC is based. The second one, namely organizational procedures, distinguishes the individual and collective procedural sequence of cognitive and physical acts related to the conception, design and production of goods. The final one, that is the input-output relation, pertains to the mainstream analysis of the relationship between a list of production inputs and some measures of output. The focus on knowledge creation requires a special attention on the first two “primitive” analytical levels while developing a systematic relationship with the third level, that is the ”derived” production process (Dosi and Grazzi, 2006).

The second challenge is to depict a broader perspective of the relationship between TC and growth by capturing the insights gained by related streams of literature. This second step would necessarily require an outstanding analytical and methodological effort because of the higher degree of complexity which characterises this framework. Such complexity refers to the departure from some simplistic assumptions towards a more accurate theoretical approach aimed at generating virtuous feedbacks with the growing available empirical evidence. A striking example refers to the literature on technology diffusion and the need to modify new growth theories in order to obtain

\[\text{Indeed, these competing models of TC and growth appear more as elements of a not-yet-invented general theory of the sources of TC rather than self-contained theoretical frameworks (Ruttan, 1996).}\]

\[\text{This pertains, for instance, to the representation of the relationship between TC and growth in terms of steady state vs. disequilibrium analysis and the different hypotheses behind an agent’s rationality (Section 4.1).}\]

\[\text{An example of such representation is given by Acemoglu and Robinson (2000).}\]
the well-established empirical S-shaped pattern of technology diffusion. Indeed, this evidence is neglected nowadays. Moreover, it is important to relate the diffusion process to the rate of TC and economic growth. This means a wider approach to the issue of technology diffusion rather than a mere focus on its measurement (Soete and Turner, 1984). Indeed, diffusion studies can contribute to the literature on knowledge investments, spillovers, externalities and they can provide a useful framework for the literature on embodied TC. This line of research represents another possible departure for an extension of endogenous growth models (see, for instance, Chari and Hopenhayn, 1991). In fact, new growth theory (Romer, 1990; Aghion and Howitt, 1998) considers only a uniform vintage of capital and, thus, it neglects a crucial source of heterogeneity in capital productivity between firms, sectors and countries (Section 3.2). Indeed, the widespread diffusion of ICTs, their fast replacement rate, and their growing intellectual content, pose new important analytical and methodological problems for a correct representation and measure of TC evolution.

Future interesting paths of research on the relationship between TC and growth appear also further analyses of the different effects of basic and applied research (Pavitt, 1991 and Brooks, 1994), the relationship of growth and TC with business cycles (Farmer and Guo, 1994) and the functioning of labour markets (Galor and Tsiddon, 1997; Galor and Moav, 2000; Lloyd-Ellis, 1999).

To sum up, every step toward further analytical enlargements of the relationship between TC and economic growth has to deal, on one side, with a more correct representation of the working mechanisms of real markets and social actors and, on the other side, with a continuous, faster and complex evolution of TC. This requires a strong theoretical knowledge of existing literature and the ability to interpret future developments of the interactions between economic factors and TC dynamics. This survey has attempted to provide such contribution.
References


Vol. 60(No. 5), pp. 1127–1150.

Ijiri, Yuji, & Simon, Herbert A. 1974. Interpretations of Departures from the Pareto Curve Firm-Size 


nomics, Vol. 110(No. 4), pp. 1127–1170.

Jaffe, Adam B. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms’ 
984–1001.


79(No. 5), pp. 957–970.


Jaffe, Adam B., Trajtenberg, Manuel, & Henderson, Rebecca. 1993. Geographic Localization of 
Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics, 
Vol. 108(No. 3), pp. 577–598.


