Innovation Heterogeneity, Schumpeterian Growth and Evolutionary Theorizing

Eduardo Pol

and

Peter Carroll

WP 04-21

December 2004
Innovation Heterogeneity, Schumpeterian Growth and Evolutionary Theorizing

Eduardo Pol*
School of Economics and Information Systems
University of Wollongong

and

Peter Carroll
University of Tasmania

Abstract

Schumpeterian growth models revolve around two tacit assumptions that are at odds with the empirical evidence, namely: all innovations are equally important for economic growth (equipollent innovation) and all innovations occur in one sector only (confined innovation). The present paper shows that it is possible to dispose of both implicit assumptions by disaggregating the ‘ideas production function’ without altering the gist of the theoretical framework. The paper refers briefly to the concepts of macro and microinventions, and introduces the concept of ‘innovatory discontinuity’. The extended theoretical framework developed here throws light on the ongoing controversy between neoclassical and evolutionary theorizing.

Key words: Innovation heterogeneity, ideas production function, scale effects problem, innovatory discontinuity, neoclassical and evolutionary theorizing

JEL classification numbers: O31, O32

* Corresponding author. Postal Address: School of Economics and Information Systems, University of Wollongong, Wollongong NSW 2522, Australia. Email: epol@uow.edu.au
I. Introduction

It is generally agreed that the work of Joseph A. Schumpeter provided much of the basis for the following axiom: to understand economic evolution it is necessary to think carefully about business innovation. In thinking about business innovation one of its characterizing features is disparateness, the lack of homogeneity. There is an intricate ‘innovation jungle’ in the real world because both the economic impact of novelty and the opportunities for developing new ideas are not uniformly distributed across sectors. We refer to this phenomenon as innovation heterogeneity.

Innovation heterogeneity has at least two empirical dimensions. One, innovations vary in terms of the magnitude of their economic impact, some having widespread effects and others being of very limited scope. Two, the fact that economic sectors vary according to sources and rates of innovation, that is, innovation is a process that occurs differently across sectors. The historical origins of these empirical understandings can be found in the papers by Simon Kuznets (1929) and Keith Pavitt (1984), respectively. These papers are excellent examples of Marshall’s dictum: “It is the business of economics, as almost every other science, to collect facts, to arrange and interpret them, and to draw inferences from them.”

It is also generally agreed that the ultimate end of the Schumpeterian vision of economic evolution is to understand fully the rules that govern a profit-oriented, market-guided economy where the increase in the standard of living of its residents is primarily based on the production of profitable new ideas. For lack of a better term we call this special kind of economy a creative economy.¹

There are two central approaches to the analysis of a creative economy: the evolutionary approach as originated in the book by Nelson and Winter (1982) and the neoclassical approach emerging from Romer (1990b). Although these approaches discuss similar economic issues concerning the creative economy, they are irreconcilable. In essence, the neoclassical approach deals with general equilibrium models of economic growth where technological change is endogenously determined by the optimizing behaviour of firms and consumers. The essential distinguishing feature of all of these Schumpeterian growth models

¹ It follows at once from this definition that a creative economy is an ideas-driven economy.
is the existence of an ‘ideas production function’ for the economy as a whole that makes endogenous the production of innovations.

The Schumpeterian growth models revolve around two tacit assumptions that are at odds with the empirical evidence, namely: first, all innovations are equally important for economic growth (e.g. the economic impact of a new satellite technology is indistinguishable from that of a new can opener), and second, all innovations occur in one sector only (the ideas-producing sector). These assumptions will be referred to as equipollent innovation and confined innovation, respectively.

The present paper shows that it is possible to relax the assumptions of equipollent innovation and confined innovation by disaggregating the ‘ideas production function’ without altering the gist of the Schumpeterian growth models. Specifically, this paper considers two stylized ideas-producing sectors, one (the enabling sector) generates innovations that have magnifying effects on the other sector (the recipient sector), but the recipient sector has no perceptible influence on the new ideas emerging from the enabling sector. Furthermore, the proposed disaggregation of the ideas production function allows for the elimination of the scale effects on the economy’s growth rate.

The paper also uses the concepts of ‘macroinvention’ and ‘microinvention’ proposed by Mokyr (1990) to introduce the concept of ‘innovatory discontinuity’. This extension of the Schumpeterian growth models throws light on the ongoing controversy of neoclassical versus evolutionary theorizing.

The remainder of the paper is organized as follows. Section II briefly outlines the conceptual framework of the Schumpeterian growth models with particular regard to the ideas production function. Section III relaxes the assumptions of equipollent and confined innovations. Section IV introduces the concept of innovatory discontinuity. In Section V we make contact with evolutionary theorizing. Finally, Section VI summarizes the conclusions.
II. Schumpeterian Growth Models

The centrality of technological innovation in economic growth has been clearly recognized by many economists ranging from Adam Smith (1776) to Abramovitz (1952) and Solow (1956). However, it is only in the recent past that technological change has been mathematically treated as an endogenous variable in a general equilibrium model capturing important aspects of the Schumpeterian vision of economic evolution. This line of research was initiated by Paul M. Romer -somewhat roughly in Romer (1990a) (1990b)- and provoked an explosion of articles on innovation and economic growth.

II.1 Romer Model: Verbal Description

Romer (1990b) developed the first general equilibrium model of a creative economy.²

The most important achievement of his model is the integration of the following four insights into a coherent conceptual framework:

- Schumpeter’s insight (New ideas): The act of innovation consists of reconfiguring old ideas in new ways to produce new ideas. Schumpeter (1934, p. 68).
- Schmookler’s insight (Profit motive and new ideas): Innovation is essentially an economic phenomenon, or at least explicable in economic terms. Schmookler (1966, p. 208).
- Romer’s insight (Nonrivalry and new ideas): The existence of intangible inputs renders increasing returns inevitable. Romer (1990a, p. 97).

The basic concepts in Romer's conceptual framework are the 'profitable new idea' exhibiting two attributes (nonrivalry and partial excludability) and ‘human capital’ with the properties of a private good. Ideas with economic value and human capital are driving force or the wheels of the creative economy. These two concepts are related to each other through an ideas production function involving two explanatory variables (the number of researchers and the stock of ideas available to these 'ideas workers') and a single dependent variable, defined as the rate of new ideas creation.

² We will refer to this contribution as Romer model.
Romer’s model also indicates that innovations only occur in the ideas-producing sector completely described by the ideas production function. There is a one-to-one correspondence between innovations and profitable new ideas. Innovations are largely stimulated by the profit motive (Schmookler’s insight) and the corresponding new ideas are at least partially excludable due to the existence of intellectual property rights (Nelson’s insight). Consequently, private investment in innovation occurs in an imperfectly competitive environment. The logic of the existence of increasing returns (Romer’s insight) is as follows. A new idea is nonrival in the sense that its use in one activity does not prevent its use elsewhere. Moreover, any new idea needs only to be created once, so that an innovation only entails fixed costs, given by the one-time costs of creating the idea. Consequently, a creative economy displays increasing returns to scale.\(^3\)

\textit{II.2 Ideas Production Function}

It is a basic premise of the new generation of formal growth models that the ideas production function reflects the innovation process in a creative economy. The general expression of the (aggregate) ideas production function (briefly, IPF) can be written as the differential equation

\[ \dot{A}(t) = F(L_A, A), \quad (1) \]

where \( \dot{A}(t) \) represents the rate of new ideas creation at time \( t \), and \( L_A \) and \( A \) denote, respectively, the number of researchers and the stock of ideas. Romer (1990b) was the first economist to make the ideas production function explicit and concrete.\(^4\)

A pictorial description of the IPF is the familiar source-target picture shown in Fig.1. The IPF turns out to be a mapping from a point in a two-dimensional space into a point in a one-dimensional space: in correspondence with each ordered pair \([L_A(t), A(t)]\) there is one and only one instantaneous rate of new ideas creation \( \dot{A}(t) \).

\(^3\) Intuitively, an increase of 1% in all inputs results in an increase in output by more than 1% because, by definition, non-rival inputs can be used over and over again simultaneously by many people.

\(^4\) He introduced a special version of this function where the rate of new ideas creation is a linear homogenous function of the number of the research workers and the stock of existing ideas. His special case automatically implies an exponential growth of the number of ideas. Romer (1990b, p. S83).
An immediate implication of the existence of the IPF should be emphasized. The specification of the right hand side of (1) would allow us to obtain a function $A(t)$ that reconstructs the past and predicts the future number of ideas in the creative economy. Indeed, a particular solution to the differential equation (1) would give a whole function $A(t)$ describing the state of knowledge at any particular time $t$. It is assumed that this function condenses or summarizes the existing state of technology.

II.3 Scale Effects

One of the central implications of the Romer model can be paraphrased as follows: if the level of resources devoted to innovation is doubled, then the growth rate of output should also double. In general, many Schumpeterian growth models exhibit the scale effects prediction: the standard of living in larger economies, which devote greater resources to innovation, should grow faster. Romer (1996) views scale effects as an important outcome of his model.

If we equate the amount of resources allocated to innovation with the level of R&D effort then our formal model exhibits the scale effects prediction if the model displays at least one of the following ceteris paribus observable consequences:

$P_1$: R&D effort and per capita income change in the same proportion and direction, or

$P_2$: the economy's growth rate is unitary elastic with respect to the level of R&D effort.

The influential paper by Charles I. Jones (1995a) forced economists to consider whether the scale effects prediction emerging from the Schumpeterian growth models is really a sensible implication. To be more precise, the so-called scale effects problem arises because there is no clear empirical evidence supporting the veracity of predictions $P_1$ and $P_2$. Dinopolous and Thompson (1999, esp. pp. 160-168). Three basic questions immediately suggest themselves:

---

5 Given the number of ideas workers, that is, once the form of the function $L_A(t)$ is specified, a differential equation like (1) usually has a general solution that depends on one constant that can be uniquely determined. For example, assuming that the stock of ideas is historically known at the initial time $t = 0$, say $A(0) = A_0$, the particular solution of equation (1) can be determined.
(a) What is the *ultimate* reason or explanation for the existence of scale effects? (b) What is the nature of the IPF? and (c) Can we get eliminate the concept of an ideas production function and still have a formal Schumpeterian model? The short answers are (a) the existence of these effects is inextricably linked to the existence of the aggregate IPF; (b) the IPF is a black box; and (c) the elimination of the IPF would imply a return to the Solow (1956) model.

Not surprisingly, given the above theorists have begun to construct Schumpeterian growth models that *exclude* the scale effects predictions. Indeed, in a short period of time Jones (1995a) has provoked several responses conducive to the removal of the scale effects prediction. In summary, these responses in the form of amended Schumpeterian growth models dealing with the scale effects problem formulate a *multiplicative specification* of the function IPF:

\[
\dot{A} = \alpha L A^\phi \tag{2}
\]

where \(\alpha\) is a constant of proportionality and \(\phi\) is an externality-related parameter, and they include a strategic supposition, namely: there are diminishing returns to innovative effort (in symbols, \(\phi < 0\)). The intuition behind this strategic supposition is that \(\dot{A}\) decreases with the level of knowledge because prior research has discovered the ideas that are easiest to find, making new ideas creation more difficult. Dinopoulos and Thompson (1999, p. 171).

Two points should be noticed. First, the amended Schumpeterian growth models have been able to remove prediction \(P_2\), but not prediction \(P_1\). Second, the amended models retain the two implicit assumptions of earlier models, namely: equipollent innovation and confined innovation.

III. Disaggregation of the Ideas Production Function

The importance of distinguishing innovations in terms of the magnitude of their economic impact goes back to Kuznets (1929). He formulated an innovation law applicable only to major innovations that can be condensed as follows: the introduction of a major technological

---

6 This point is forcibly made by Jones (1999, p. 143).
innovation in a given sector leads to a phase of rapid sectoral growth and gradually generates a set of forces leading to a deceleration in the rate of growth of the sector in question.\(^7\)

As Pavitt (1984) has shown, sectors differ in important innovation aspects. His appreciative theory of the creative economy emphasizes the flows of innovation between sectors and that technological change occurs differently across industries. There are several ideas sectors operating in a creative economy, namely: science-based sectors developing major technological innovations, specialist supplier sectors existing in symbiosis with scale-intensive sectors where firms develop mainly minor innovations, and supplier-dominated sectors developing only minor technological innovations.

A glance at Fig.1 shows that the formulation of the innovation process implied by the IPF is extremely condensed, not including many important aspects such as possible distinctions between ‘big’ ideas and ‘small’ ideas, sectoral patterns of innovation, etc. The relaxation of the assumptions of equipollent innovation and confined innovation also requires that we enter the black box of the IPF and develop an explicit model of innovation production which incorporates the following two insights:

the Kuznets Insight: Innovations differ in terms of their economic impact.

the Pavitt insight: Innovation occurs differently across sectors and there is a variety of innovation linkages between enabling and recipient sectors.

Consider two ideas-producing sectors, Sector 1 (enabling sector) and Sector 2 (recipient sector). At any time \(t\) the state of technology \(A(t)\) is decomposed into ideas generated in the enabling sector \(A_1(t)\) plus ideas originated in the recipient sector \(A_2(t)\)

\[
A(t) = A_1(t) + A_2(t) \tag{3}
\]

We assume that new ideas emerging from the enabling sector have a multiplier effect in the recipient sector, but the new ideas created in the recipient sector do not have a perceptible

\(^7\) Some 40 years after the formulation of the law of retardation of sectoral growth, Kuznets wrote an illuminating paper on the impact of major innovations Kuznets (1972). For a detailed analysis of the Kuznets law on innovation, see Pol and Carroll (2004).
influence on the generation of new ideas in Sector 1. In symbols, the ideas production functions for Sectors 1 and 2 can be written, respectively, as

\[
\dot{A}_1(t) = L_1(t)[A_1(t)]^\mu
\]
\[
\dot{A}_2(t) = L_2(t)A_1(t)[A_2(t)]^\nu,
\]

where \(L_1(t)\) and \(L_2(t)\) denote the amount of labour allocated to producing ideas in Sectors 1 and 2, respectively, and \(\mu\) and \(\nu\) are the externality parameters in the enabling and recipient sector, respectively. For lack of a better expression we call the system of differential equations (4)-(5) together with the identity (3) the \textit{innovation regime}.

Following a line of reasoning identical to Romer (1990b), it can be easily shown that an equilibrium of the extended model gives the paths for prices and quantities corresponding to a pre-assigned set of parameters such as the stock of human capital and final output elasticities, and that the market mechanism does not lead the creative economy to an optimum due to the existence of intellectual property rights. The strategy concerning this proof consists of integrating the system of separable differential equations (4)-(5) to obtain the state of technology \(A(t)\). As to the scale effects problem, it can be seen that the assumption of diminishing returns to innovative effort in each ideas-producing sector (that is, \(\mu < 0\) and \(\nu < 0\)) is sufficient to eliminate prediction \(P_2\). Regrettably, prediction \(P_1\) cannot be removed through the suggested disaggregation of the ideas production function.

To sum up, the preceding disaggregation of the ideas production function incorporates a stylized notion of innovation heterogeneity without altering the gist of the Schumpeterian growth models.

\textbf{IV. Macroinventions and Innovatory Discontinuities}

Joel Mokyr (1990) suggests that economic evolution can be seen as a multidimensional object that is characterized by 'macroinventions' (radical new ideas involving major technological breakthroughs) and 'microinventions' (small improvements to existing techniques). He dissents from the view that technological change is devoid of discontinuities (Mokyr, 1990, esp. pp. 12-13).
The emphasis on ‘innovatory discontinuities’ harks back to Schumpeter himself. He was quite explicit about the discontinuous nature of macroinventions, although he did not use the term macroinvention in his writings. His analysis in Business Cycles was intended to apply only to radical innovations of a kind that implied a significant shift of the state of technology. In Schumpeter’s own words:

(...)

We shall impose a restriction on our concept of innovation and henceforth understand by an innovation a change in some production function which is of the first and not of the second or a still higher order of magnitude. A number of the propositions which will be read in this book are true only of innovation in this restricted sense.

Schumpeter (1939, p. 94) [Italics in original]

One of the distinguishing features of an innovation is that it can always be understood ex post, but it can never be fully understood ex ante applying the ordinary rules of inference to the existing facts. Innovation is by definition an uncertain phenomenon. However, for analytical purposes the difference between a macroinvention and a microinvention is that macroinventions are shrouded in Knightian uncertainty (sensu stricto uncertainty) while microinventions are susceptible of calculable uncertainty (risk in Knight's sense). This approximation is in line with the history of technological innovation (Mokyr, 1990, esp. p. 295).

In the Schumpeterian growth models, the state of technology A(t) presupposes that all macroinventions have already occurred and that technological change consists of a continuous sequence of microinventions. By definition, a macroinvention (e.g. the invention of the electricity) substantially alters the prevailing state of technology, and thereby provoke a change in the current innovation regime. After a macroinvention has occurred the new innovation regime can be mathematically described as

\[
\dot{B}_1(t) = L_1(t)[B_1(t)]^\theta, \quad (6)
\]

\[
\dot{B}_2(t) = L_2(t)[B_2(t)]^\phi [B_1(t)]^\theta, \quad (7)
\]
where $\theta$ and $\rho$ are the externalities parameters corresponding to the new situation, and the state of technology is now

$$B(t) = B_1(t) + B_2(t)$$ (8)

In brief, when a macroinvention occurs the state of technology changes from $A(t)$ to $B(t)$.

Macroinventions can be thought of as *random innovation shocks* affecting the whole creative economy. An innovation shock entails a *discontinuity* in the following sense: the state of technology changes from $B(t)$ to $A(t)$. Specifically, an *innovatory discontinuity* is said to occur when a macroinvention provokes a selective replacement of the state of technology. An innovatory discontinuity is not necessarily a ‘jump’ discontinuity but rather, and perhaps more typically, a gradual change from one state of technology to another.

The empirical intuition behind the notion of innovatory discontinuity can be illustrated by using the examples of the electricity and IT eras. Electrification arrived in the 1890s (the start-up of the electricity era is often taken as the construction of the first hydro-electric facility at Niagara Falls in 1894) and from the viewpoint of technology adoption attained a plateau in 1929. The IT era started in 1971 (when Intel’s invention of the key component of the personal computer occurred (namely, the “4004 computer chip”) and still underway.

V. Evolutionary Versus Neoclassical Theorizing

According to evolutionary scholars, the new generation of formal growth models represents a desirable convergence of formal theory with appreciative theory. They combine important aspects of reality (such as innovation, imperfect competition, proprietary aspects of technology, and increasing returns to scale) within a general equilibrium framework. Nelson (1994, p. 309)

However, evolutionary theorists believe that the neoclassical formulation is inconsistent with the Schumpeterian argument that a creative economy had to be understood as a process inextricably linked to *disequilibrium*. Specifically, the evolutionary approach entails the throwing away of both the equilibrium and optimizing notions that constitute the unifying

It should be clear that although the central focus of both neoclassical and evolutionary theorizing is economic evolution, these approaches concentrate on two different meanings of the word ‘evolution’: gradual predictable change and erratic change based on mutation and selection, respectively. Indeed, while the neoclassical vision indicates that economic growth and technological change are predictable and continuous, the evolutionary vision stresses the unpredictability of macroinventions and rejects the premise that technological change is devoid of discontinuities.

The extended analytical framework developed in the previous section throws light on the role of the two senses of the term evolution. If we assume that the creative economy is operating with the state of technology represented by the function \( A(t) \) and that a macroinvention occurs at a particular point in time \( t^* \) (see Fig.2), then how does the creative economy move from one state of technology to another? There is no obvious answer. The propagation mechanisms are difficult, if not impossible to decipher ex ante because technologies never move in a predictable fashion. For example, formerly unconnected technologies (such as lasers and fibre optics) may turn out to be complementary.

**FIGURE 2 HERE**

After the occurrence of a macroinvention there is a transitional dynamics converging to the new state of technology \( B(t) \). There exists an interval (not just a point) of discontinuity \( t^{**} - t^* \) where the transition from the old to the new state of technology takes place. What is involved in this transitional dynamics is an extensive process of technological cross-pollination, redesign, modification, and innumerable small improvements occurring after the introduction of a macroinvention. This means that the convergence to the new state of technology \( B(t) \) will take a lengthy period of time (technologies move slowly from the first macroinvention) and the transitional dynamics is, at least in part, intrinsically intractable.
Figure 2 shows the irreconcilability of the neoclassical and evolutionary approaches. The neoclassical approach assumes gradual predictable evolution as represented by the states of technology $A(t)$ or $B(t)$. The Schumpeterian growth models concentrate on evolution in the first sense and ignore the transitional dynamics. Evolutionary theorists tend to focus on stochastic evolution characterized by Knightian uncertainty and disequilibrium conditions. They concentrate on the second sense of the term evolution and reject the use of optimizing theories.

If we accept that technological change is path dependent (somewhat roughly, an evolutionary process taking place under uncertainty in the strict sense) one implication is that it is virtually impossible to theorize about technological developments. In particular, the production of new ideas seems to depend on the random history of the creative economy. There can be no well-defined ideas production function. Evolutionary theorists do not agree with the view the existence of uncertainty in the Knight’s sense implies that ‘anything goes’. The recent work of Cristiano Antonelli (2003), for example, builds a bridge between the economics of innovation and the economics of technological change and provides an enlarged appreciative framework for analyzing the implications of the introduction of a new technological system in the global economy. The conceptual model elaborated in his book separates *internal* path-dependence from *external* path-dependence in order to explain the interconnections between technological change and structural change. Unexpected changes and disequilibrium conditions in the market place act as an impulse for the creation of new ideas.

**VI. Summary**

The past decade has been marked by a number of important developments in the analysis of a creative economy. Schumpeterian growth models have articulated four insights in a general equilibrium context emerging from optimizing behaviour, namely: that innovation has a recombinant nature (Schumpeter insight), that innovation is pursued for gain (Schmookler insight), that new ideas are at least partially excludable (Nelson insight), and that innovation generates increasing returns to scale (Romer insight). In the simplified world of the Schumpeterian growth models technological innovations come from an *ideas-producing*

---

8 An earlier empirical study dealing with the life-cycle of major technological innovations is due to Kuznets
sector which operates according to the ideas production function. Furthermore, technological change is viewed as a cumulation of small, individually minor innovations. The assumption that there exists an aggregate ideas production function is the sine qua non of the Schumpeterian growth models. Although the analytical meaning of such a production function is simple and clear, troublesome questions arise when one wants to introduce the notion of innovation heterogeneity into these formal models.

Our analysis demonstrates that it is possible to relax two simplifying assumptions underlying the Schumpeterian growth models (termed here equipollent innovation and confined innovation) without altering the gist of these formal models. Specifically, we have incorporated two additional insights: one, that innovations differ in terms of their economic impact (Kuznets insight) and two, that there are linkages between enabling sectors and recipient sectors (Pavitt’s insight), evident when disaggregating the ideas production function.

Early Schumpeterian growth models led to the logical conclusion that growth rates increase with economy size (scale effects prediction). These scale effects can be easily eliminated from our extended framework using the assumption of diminishing returns to innovative effort in both enabling and recipient sectors.

A large proportion of the total growth in productivity takes the form of microinventions. However, there have been large and spectacular changes in technology leading to economically significant changes as illustrated by the electricity and IT eras. Thus, the empirical evidence supports the view that there can be nonincremental technological changes and drastic technological shocks to the economy. Our concept of innovatory discontinuity it is intended to these events.

It is generally recognized that the neoclassical and the evolutionary approaches to the study of technological change are irreconcilable. Neoclassical and evolutionary theorizing use the term evolution in different senses. In the context of the Schumpeterian growth models evolution means gradual predictable change devoid of discontinuities. For the evolutionary theorists, the word evolution refers to the unfolding of a fundamentally unknown future where there is change based on mutation and selection and discontinuities can and do happen.

(1929).
Evolutionary scholars believe that the economics profession will ultimately be driven to adopt their non-equilibrium approach, if economists attach high priority to characterizing and modelling unforeseen economic change induced by technological shocks. Mainstream economists seem to believe that the style of modelling used by the neoclassical economists is appropriate because the equilibrium concept is flexible enough to encompass a time path along which the salient variables change in a predictable manner (moving equilibrium).
References


Figure 1
Figure 2