NBER WORKING PAPER SERIES

SUBSTITUTION AND COMPLEMENTARITY IN ENDOGENOUS INNOVATION

Alwyn Young

Working Paper No. 4256

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1993

Sloan School of Management, MIT. I am grateful to Olivier Blanchard, Elhanan Helpman, Paul Romer, Julio Rotemberg, Andrei Shleifer and two anonymous referees for helpful comments. This paper is part of NBER's research program in Growth. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #4256
January 1993

SUBSTITUTION AND COMPLEMENTARITY IN ENDOGENOUS INNOVATION

ABSTRACT

The influence of Schumpeter's notion of "creative destruction" may have led to an overemphasis on substitution between technologies in recent models of endogenous innovation. Historical examples of technological change suggest that new technologies may just as frequently complement older technologies, creating, rather than destroying, rents. Acknowledgement of the potential for both substitution and complementarity amongst inventions allows for a much richer characterization of the growth process, creating the possibility of threshold effects and multiple equilibria, and bringing to the forefront the important role played by the expectations of inventive entrepreneurs.

Alwyn Young
Sloan School of Management
MIT
Room E52-446
50 Memorial Drive
Cambridge, MA 02139
and NBER

I INTRODUCTION

Recent models of endogenous innovation, e.g. Aghion & Howitt [1992], Grossman and Helpman [1991a] and Romer [1990], embody the Schumpeterian view of innovation as a process of creative destruction, in which new products "steal" the rents accruing to older technologies by substituting for these in either consumption or production. Without any doubt, this focus on the potential substitutability of different technologies accords well with many popular examples of technical change, perhaps the most influential of which has been the development of computer microprocessors and memory chips, where each successive generation of new and improved inputs has rapidly replaced the previous generation in almost all applications. Yet, as noted by Rosenberg [1976], the Schumpeterian emphasis on substitution ignores the equally important complementarity that exists between many technologies, where, oftentimes, new inventions create, rather than destroy, rents for older technologies. Most new technologies, at the moment of their invention, are decidedly inferior to existing products and processes. As such they pose little immediate threat to existing technologies, finding only limited applications where they, more often than not, merely serve to enhance the efficiency and profitability of older systems. Typically, a sequence of additional innovations serve to "round out" new technologies, allowing these to ultimately replace, en masse, older systems. In essence, many new technologies pass through a life cycle, in which they initially complement older technologies, but, subsequently, as additional inventions which are complementary to them appear, successfully substitute for the older techniques.

Some historical examples might serve to illustrate the life cycle effects I have in mind. The steam engine, as invented by Newcomen in the early 18th century, was little more than a crude piston, whose vertical motion posed no immediate threat whatsoever to the dominance of the water wheel in the provision of the rotary force essential to the operation of most machinery.

Watt's (1765) invention of the separate condenser, which eliminated the grossly inefficient cooling and reheating of the main cylinder, and Wilkinson's (1774) invention of the boring mill, which finally provided a means of engineering a tight fit between cylinder and piston, increased the efficiency of the system by orders of magnitude. Nevertheless, the steam engine still remained little more than a water pump, whose principal applications involved pumping water out of mines and, significantly, enhancing the operations of water wheels by recycling water back up to the upper mill pond. The invention of the "sun and planet" gearing system by Murdock in 1781 finally provided a means of converting vertical motion into rotary force, allowing the widespread application of steam power on the factory floor. Even so, the replacement of the water wheel was gradual and dependent upon continued improvements in the efficiency of the steam engine. As late as 1869 fully 48.2% of all primary power in US manufacturing was still provided by water wheels and turbines, which provided a more regular speed, required less maintenance, and used a cheaper fuel than steam engines.[1]

The gradual replacement of the sailing ship by the steam ship provides another interesting example of the importance of complementary effects between technologies. In the 1820s steam ships began moving trade between England and North Sea ports. By the 1830s they were plying routes throughout the Mediterranean and on to India, while by the 1840s British steam vessels had reached the United States and South America. Despite this promising early growth, the replacement of the sailing ship by the steam ship as the dominant form of oceanic transport took the better part of the century. Steam screws required sturdy iron hulls, which, however, had a tendency to foul. Combined with the need to carry onboard large quantities of coal fuel to feed

[1] Burke (1985, ch. 6), Law, R.J. (1965), Rosenberg (1976, chs. 10 and 11).

the engines,[2] this actually made early steam ships slower, less reliable and more costly to operate than sailing ships on long voyages. Sailing ships also had lower (stationary) operating costs, which meant that they could be unloaded more slowly when in port and, during the depression of the 1870s, could afford to wait for cargo rather than lay up. Furthermore, many of the new technologies which underlay the development of the steam ship actually served to enhance the efficiency of sailing ships. Light composite hulls of iron encased in wood, steam driven winches, auxiliary steam engines, and the use of steam tugs for harbour towing, which allowed the construction of large capacity vessels, all combined to maintain the competitiveness of the sailing ship. Not until the development of the high pressure triple expansion engine in the 1880s increased engine efficiency ten-fold, did the steam ship attain a decisive margin of superiority over sailing ships in most applications. Ironically, the widespread replacement of sail by steam gave the sailing ship one last lease on life, as the low cost carrier of coal to far flung coaling stations. In 1892, after seventy years of substitution and complementarity between steam and sail, British shipyards launched their largest tonnage of new sailing ships ever.[3]

Much of the complementary development and modification of new technologies described above might most appropriately be viewed as the outcome of simple experience in their use, i.e. a serendipitous process of learning by doing in which the actors who stumble across the improvements do not necessarily acquire rents from so doing.[4] In this case, it is the balance between the incentives to engage in research, where actors seek to acquire rents, and the incen-

---

[2] At the time, England was the world's principal source of coal.

[3] Graham (1956), Harley (1971), Rosenberg (1976, ch. 11).

[4] For example, Mak and Walton [1972] emphasize the important role played by experience in improving the productivity of Mississippi steamboats during the early 19th century in what was increasingly a perfectly competitive market without excess returns (see Haites and Mak 1970).

tives to engage in production, where actors may not gain rents from generating new knowledge, that determines an economy's growth rate. I have explored this perspective, and the importance of a complementary balance between inventive and productive activity that it implies, in an earlier paper (Young 1992). At the same time, it must be acknowledged that much of the complementary innovation which enhances the commercial viability of new technologies is, without a doubt, the result of rent seeking inventive activity on the part of other innovators.[5] This creates a situation in which the inventive activity of rent seekers depends positively on the inventive activity of other rent seekers, under which circumstances the self-fulfilling beliefs of inventors may become the most important determinant of the economy's growth rate. This paper explores this perspective.

Section II below presents a simple model of invention with substitution and complementarity. The model consists of a two tiered production structure in which intermediate inputs are used in the production of final goods, each of which makes use of a subset of all available inputs. The invention of new intermediates allows the production of new final goods. The dissipation of consumer expenditure across this enlarged set of final goods and, by extension, intermediate inputs lowers the profits accruing to each intermediate input patent holder. In addition, a subset of existing final goods producers will make use of the new intermediate input technologies, which, for a given pattern of consumer expenditure, shifts producer demand away from existing inputs. These expenditure switching effects make inventions substitutes, as a rise in the number of technologies lowers the profit flow accruing to each patent holder. At the same time, however, the production of new final goods (allowed by the invention of new intermediates) creates

[5] The contemporary efforts of computer hardware producers, such as Next, to entice software companies into the costly development of complementary software is a nice example of complementarities between rent seekers.

new applications for existing inputs, which raises their profitability. In this sense, inventions are complementary.[6]

When an intermediate input is initially invented it lies at the forefront of industrial technology and faces a relatively small market, being of use in the production of only comparatively advanced final goods. Consequently, the market creation effect dominates and further invention is complementary to it. Once a technology has matured, that is, faces a large and developed market, the expenditure switching effect becomes dominant, and further invention lowers the flow of profits to its patent holder. Thus, inventions follow a life cycle, finding further invention complementary early in life, but substitutable in maturity. Although a model with a single linear dimension cannot hope to capture all of the complex interrelations present in the history of technological change, I believed that this stylized structure does provide some insight into the evolving tension between substitution and complementarity characteristic of so many technologies.

Introducing complementarities between the actions of economic actors immediately creates the possibility of multiple equilibria. Section II shows that the model of this paper may have as many as three steady state equilibria: complete stagnation, a low growth equilibrium in which complementarity between inventions is dominant, and a high growth equilibrium in which substitution is dominant. Traditional models of endogenous innovation correspond to a special case of the model, where only one (substitution dominated) steady state equilibrium exists. The steady state growth rates of the model are also discontinuous in its parameters. In particular, a threshold effect exists in which small parameter changes or policy actions may discontinuously

---

[6] The reader will note that I am using a terminology analogous to that of the game theoretic literature in associating the terms complementarity and substitutability with the effects of further invention on the payoff (profitability) of existing firms. The narrower microeconomic definition of complementarity and substituability, with its emphasis on cross derivatives in production and utility functions, is an important source of payoff complementarities, but not the only one.

enlarge the opportunity set of a stagnant economy by creating an additional, unique, high growth equilibrium.

In standard substitution-dominated models of invention the fact that all out of steady state paths diverge explosively from the steady state is customarily used to justify an immediate movement to the steady state. An analysis of perfect foresight dynamics in section II shows that the substitution dominated steady state is, when it exists, indeed unstable. Many of these divergent paths, however, converge to the complementary steady state, which, when it also exists, is stable. This, however, does not exhaust the potential dynamics of the model. The simple model of section II makes use of preferences which exhibit an extraordinary preference for variety, which, as it so happens, in general equilibrium serve to suppress some of the life cycle effects in the model, greatly simplifying its analytics. In section III I show how the use of standard preferences to close the model enormously enlarges the range of potential equilibria, including allowing for conditions under which *both* the complementarity and the substitution dominated steady state are stable. In general, the range of dynamic behavior allowable, and its sensititivity to changes in the preference structure used to close the model, is remarkable. Section IV concludes.

II A SIMPLE MODEL OF INVENTION WITH SUBSTITUTION AND COMPLEMENTARITY

(A) Model Structure

Consider an economy with only one factor of production, labour, which is in fixed total supply L. Labour can be used to produce or invent intermediate inputs, which are ordered along the real line and indexed by v. At any point in time, the economy only knows how to produce a finite set of intermediate inputs, those laying in [0,N(t)], with $N(t)^{-1}$ units of labour required to produce one unit of any input v.[7] New inputs are invented in competitive research labs, with the aggregate rate of invention linear in the total amount of labour devoted to research:

$$(1) \quad \dot{N}(t) = N(t)L_R(t)/a_R$$

Research firms finance their R&D efforts by issuing shares, which are traded in an asset market. Successful innovators are rewarded with an infinitely lived patent and compete monopolistically with all other patent holders, distributing any profits to their shareholders. There is free entry into research.

Final goods are produced by perfectly competitive firms, which combine intermediate inputs in a CES production function, with the output of final good s given by:

$$(2) \quad Q(s) = \left[\int_{sB}^{\mathrm{Min}[s\Theta, N(t)]} x(v,s)^{\alpha}\, dv \right]^{\frac{1}{\alpha}} \qquad \Theta \geq 1,\ 1 > B \geq 0$$

The invention of an intermediate input s creates, as an unappropriated by-product, the knowledge of how to produce final good s, which is initially produced by combining all inputs in $[sB, s]$. As further, more advanced, intermediates are invented, these too may be used in the production of

[7] This is a technical assumption, necessary to keep (logarithmic) utility growing in the steady state. It is dropped in section III, which uses CES preferences.

good s, but only up to some finite input $s\Theta$. Beyond $s\Theta$, the technical differences between industry s and newly invented inputs are so large that these inputs cannot contribute to the production of good s.

Readers familiar with the models of Romer (1987, 1990) emphasizing gains from increasing specialization will quickly see that the production function (2) above simply modifies Romer's technology to allow for input heterogeneity. In Romer's model all inputs (no matter when invented) can be used symmetrically in all existing industries.[8] If one believes that inputs tend to be invented in an explicit (heterogeneous) technological sequence, with the technological difference between any two inputs growing with the number of inventions that lie between them, then it seems reasonable to assume that a final good s, associated with input s, should only use inputs in a technological neighbourhood of input s. Having said this, however, I must note that to remain tractable the model must retain some measure of symmetry. Thus, I shall assume that $B = 0$, i.e. that new goods can always make use of all inputs in existence at the time of their invention.

The assumption of input heterogeneity immediately introduces a tension between complementarity and substitution in the process of invention. When input N(t) is invented at time t, only final goods producers in $[N(t)/\Theta, N(t)]$ make use of that input. The invention of further inputs, along with the associated knowledge of how to produce more advanced final goods, will expand the market for input N(t), raising its profitability. In this sense, inventions are complementary. At the same time, however, further invention expands the range of usable inputs available to firms in $(N(t)/\Theta, N(t)]$, which, for a given level of consumer expenditure on each product,

[8] Actually, Romer's model only has one final good which uses all N inputs. This formulation, however, is technically equivalent to one in which there are N final goods, each of which uses all N inputs in a symmetric fashion.

lowers their demand for all existing inputs in $[0, N(t)]$. In addition, with the appearance of new final goods aggregate consumer expenditure is now spread across a wider variety of final goods and, by extension, intermediate inputs. This dissipation of both consumer and producer expenditure tends to lower the profits accruing to all input producers (including firm N(t)). In this sense, as in Romer's model, inventions are substitutes. It is this tension between substitution and complementary that creates the multiple equilibria which are inherent to this model.

To close the model, we turn to consumer preferences. I shall assume that consumer preferences are additively separable across time:

$$(3) \quad U = \int_t^{\infty} e^{-\rho(\tau - t)} V(\tau)\, d\tau$$

with V(t), instantaneous utility at time t, given by:[9]

---

[9] The assumption, presented earlier above, that $N(t)^{-1}$ units of labour are required to produce one unit of any intermediate at time t was introduced to ensure that, given these logarithmic preferences, utility does not decline as the number of available varieties rises (which would happen if output per variety fell as the effective labour supply were held constant). Given the CES structure of final producer demand, intermediate input firm profits will be independent of their costs of production. Hence, this is a purely technical assumption, and is irrelevant to the mechanics of the model (although not to its normative implications).

Technically, the assumption of declining marginal cost could be justified by assuming that the invention of each input v leads to the discovery of a new sub-input u (over which no one holds a patent), with each intermediate input being produced by symmetrically combining all sub-inputs [0,N(t)]:

$$x(v) = \left[\int_0^{N(t)} q(u)^{\frac{1}{2}}\, du \right]^2$$

Given the symmetry, and assuming that one unit of labour produces one unit of each sub-input u, if a total of l units of labour are used in the production of input v, then output is given by:

$$x(v) = \left[\int_0^{N(t)} [l/N(t)]^{\frac{1}{2}}\, du \right]^2 = lN(t)$$

Thus, in these circumstances, $N(t)^{-1}$ units of labour are required to produce one unit of any input v.

(4) $V(t) = \int_0^{N(t)} \ln[C(s,t)]\, ds$

Clearly, consumers with these preferences have an exceedingly strong preference for variety, in that a failure to consume any set of goods of positive measure leads to infinitely negative utility. In general equilibrium, a change in the rate of invention will lead to a dramatic reduction in the rate of interest, as consumers try to shift consumption into the future. As it so happens, this interest rate effect exactly cancels the substitution effect due to the dissipation of consumer expenditure across a broader variety of products, which greatly simplifies the analysis of the model's dynamics. Section III of this paper shows hows the use of more standard constant elasticity preferences leads to an extraordinary expansion of the potential set of equilibrium dynamic paths, without, however, changing much of the intuition derived from the simple model presented here.

(B) General Equilibrium

In this section I review the relations which determine the intertemporal equilibrium of the model. Since the consumer's utility function is time separable, the consumer's problem can be broken down into one of maximizing instantaneous utility, V(t), subject to instantaneous expenditure, e(t), and then picking a time path of expenditure which maximizes total intertemporal utility. As is well known, logarithmic utility leads the consumer to distribute her expenditure evenly across all N(t) final goods. Thus, instantaneous expenditure on any final good s is given by $e(s,t) = \frac{e(t)}{N(t)}$. Total intertemporal utility is then given by:

(5) $$U = \int_t^{\infty} e^{-\rho(\tau-t)} \left\{ \int_0^{N(\tau)} \ln[e(\tau)/p(s,\tau)N(\tau)]\; ds \right\} d\tau$$

$$= \int_t^{\infty} e^{-\rho(\tau-t)} N(\tau) \ln[e(\tau)/N(\tau)]\; d\tau - \int_t^{\infty} e^{-\rho(\tau-t)} \left\{ \int_0^{N(\tau)} \ln[p(s,\tau)]\; ds \right\} d\tau$$

where $p(s,\tau)$ denotes the price of good s at time τ. Maximizing (5) subject to an intertemporal budget constraint leads the consumer to select a path of instantaneous expenditure which satisfies: $\hat{e}(t) = \dot{R}(t) - \rho + \hat{N}(t)$, where $\dot{R}(t)$ is the rate of interest at time t.[10]

Let E(t)=e(t)L denote aggregate expenditure. From the above, we have the equilibrium relations:

$$(6) \quad E(s,t) = \frac{E(t)}{N(t)}$$

$$(7) \quad \hat{E}(t) = \dot{R}(t) - \rho + \hat{N}(t)$$

Equation (6) illustrates how a rise in the number of inputs tends to lower the profit flow of all patent holders, as consumer expenditure is dissipated across a broader variety of goods. (7) shows how, ceteris paribus, a rise in the rate of invention leads consumers to shift expenditure into the future. In partial equilibrium, i.e. holding the interest rate constant, these two effects cancel exactly, so that a rise in the rate of invention does not change expectations of the future flow of consumer expenditure per product.[11] In general equilibrium, a rise in the rate of invention does not actually lead to an increase in the rate of growth of expenditure, but, instead, results in a compensating reduction in the real rate of interest. Thus, although firms find that a rise in the rate of invention does imply a more rapid decline in expenditure per product, and hence the current flow of profits, this is offset by a decline in the rate at which these profits are discounted. In general equilibrium, as in partial equilibrium, these two effects exactly offset

[10] I follow the conventional notation, with superscripted dots denoting derivatives with respect to time and hats the proportional rate of growth of the relevant variable.

[11] $[E(t)\hat{/}N(t)] = \hat{E}(t) - \hat{N}(t) = \dot{R}(t) - \rho$

each other.

Turning now to producers, given the CES production function for final goods, the demand for input v by producers of final good s follows the familiar CES form:[12]

$$(8)\quad x^D(v,s) = \frac{p(v)^{-\varepsilon}E(s)}{\int_0^{\min[s\Theta,N]} p(v)^{1-\varepsilon}\,dv} \qquad \varepsilon = \frac{1}{1-\alpha}$$

where p(v) denotes the price charged by the producer of input v. Given this structure of demand, intermediate input producers find it optimal to sell their product at a fixed markup over marginal cost:

$$(9)\quad p(v) = [\alpha N]^{-1} \ \forall\, v$$

where I have taken the nominal wage as the numeraire. It follows that:

$$(10)\quad x^D(v,s) = \bar{x}(s) = \frac{\alpha N E(s)}{\min[s\Theta, N]} \ \forall\, v$$

Using (9) and (10) one can easily compute the profits accruing to the producer of each input v, which equal price minus average cost times the total quantity demanded:

$$
\begin{aligned}
(11)\quad \pi(v) &= [p(v) - N^{-1}]\int_{v/\Theta}^{N} \bar{x}(s)\,ds \\
&= [(\alpha N)^{-1} - N^{-1}]\int_{v/\Theta}^{N} \frac{\alpha N E(s)}{\min[\Theta s, N]}\,ds \\
&= \frac{(1-\alpha)E}{N}\left\{\int_{v/\Theta}^{N/\Theta} (s\Theta)^{-1}\,ds + \int_{N/\Theta}^{N} N^{-1}\,ds\right\} \\
&= \frac{(1-\alpha)E}{N\Theta}[\log(N/v) + \Theta - 1]
\end{aligned}
$$

[12] I drop the notation denoting the implicit dependence of the endogenous variables on time.

It is informative to take the partial (equilibrium) derivative of the next to last line of the above expression with respect to N:

$$(12)\quad \frac{\partial\pi(v)}{\partial N} = -\frac{\pi(v)}{N} - \frac{(1-\alpha)E(\Theta-1)}{\Theta N^2} + \frac{(1-\alpha)E}{N^2}$$

The first (negative) term on the right hand side represents the loss in profitability attributable to the dissipation of consumer expenditure across a broader variety of goods[13] and is proportional to the current profit flow of firm v. Since intermediate input producers nearer the technological frontier have smaller markets, serving only final goods producers in $[v/\Theta, N]$, and hence smaller profits, this loss is decreasing in v.[14] An increase in N(t) also expands the potential range of usable inputs available to final goods producers in $(N(t)/\Theta, N(t)]$. The negative effect, for a given distribution of consumer expenditure, of this substitution by final goods producers on the profits of existing intermediate input firms is reflected in the second term on right hand side. Since the firms in $(N(t)/\Theta, N(t)]$ use all inputs in $[0, N(t)]$, this loss is independent of v. Finally,

[13] As mentioned earlier above, given the preferences used in this section in general equilibrium there will be an exactly offsetting interest rate effect. Nevertheless, I believe a partial equilibrium analysis illuminates the intuition underlying the model, particularly since, for standard CES preferences (as in section III below), the general equilibrium interest rate adjustment need not exactly offset this effect.

[14] The reader may feel uncomfortable with the fact that profitability is monotonically declining in v, implying that the most primitive input earns the greatest profit flow. This is simply a result of the fact that, for analytical simplicity, I have entered all inputs symmetrically into the final goods production functions (2). If the production functions were in some fashion weighted in favour of more advanced inputs, this effect would disappear. I have examined the sensitivity of the model in this respect by complicating the model of section III below with the assumption that v^{-1} units of labour are required to produce one effective unit of input v at all times (all inputs enter symmetrically, in effective units, into production). This makes $\pi(v)$ hump-shaped in v. Thus, mature (but not antedeluvian) industries make the highest profits. Despite the additional complexity, the results of the model are unchanged.

the third (positive) term represents the increased profits attributable to the creation of new sources of demand for all existing intermediate inputs.[15]

It is easily shown that, provided $\Theta < 2$:

$$(13)\quad \frac{\partial \pi(v)}{\partial N} > 0 \quad \forall \ v > N e^{\Theta - 2}$$

$$\frac{\partial \pi(v)}{\partial N} < 0 \quad \forall \ v < N e^{\Theta - 2}$$

Thus, provided Θ is sufficiently small, each invention will pass through a life cycle. When an input is recently invented the profit flow accruing to the holder of its patent is small, since few existing industries use that input. Consequently, further invention is complementary, with the net effect of expenditure diversion and market creation being positive. As technology progresses and a product matures (i.e. becomes comparatively backward), the expenditure diversion effect grows in relative magnitude, until further invention unambiguously lowers current profits.

The significance of Θ lies in the fact that it determines the initial market size faced by a new invention, $[N/\Theta, N]$. For Θ near one, the initial market for new inventions is small, and, hence, the overall profitability of these depends crucially upon further invention creating additional sources of demand. If Θ is large, the initial market faced by a new invention is large. Consequently the net effect of the diversion of consumer and producer expenditure brought about by further innovation is to lower current profits. Thus, whether or not complementary

[15] I took the derivative of the next to last line, rather than the last line, of (11) because the final integration combines two effects, producer substitution and market creation, into one term. The sum of these two effects is always positive, although it approaches zero as Θ becomes large. Intuitively, if consumer expenditure per product is kept constant as the total number of products rises, then aggregate expenditure is, in essence, remaining proportional to N. Although the total number of inputs is also proportional to N, the assymetry of input profits means that although average profits per input are unchanged, the addition of new inputs, with lower than average profits, raises the profitability of all existing inputs.

effects exist depends strongly upon the degree to which new inventions find large existing markets, or, equivalently, the degree to which they depend upon further invention to generate sizeable demand.

Returning now to the analysis of the general equilibrium, labour market clearing requires that the total labour used in manufacturing and invention equal the total labour supply. Clearly, total expenditure on manufactures (E) must equal total revenue, which in turn equals the markup (α^{-1}) times total labour costs. With the wage normalized to one, total labour costs equal the total labour in manufacturing. Consequently, it is easy to see that the total labour in manufacturing equals αE. If we let $g = \hat{N}$, we can see from equation (2) that the total labour used in research equals $g a_R$. Thus, labour market clearing requires that at each point in time:

$$(14) \quad L = \alpha E(t) + g(t) a_R$$

Finally, free entry into research ensures that the cost of inventing the incremental input, $a_R/N(t)$, is greater than or equal to its market value, i.e. the net present value of profits accruing to firm N(t):

$$(15) \quad \frac{a_R}{N(t)} \geq \int_t^{\infty} \pi(N(t),\tau)\, e^{-R(\tau)+R(t)}\, d\tau \qquad = \text{if } g(t) > 0$$

Since along an equilibrium path $E(\tau) = E(t) e^{R(\tau)-R(t)} e^{-\rho(\tau-t)} e^{\int_t^{\tau} g(v)\,dv}$, we may use (11) to derive:

$$(15)' \quad \frac{a_R}{N(t)} \geq \int_t^{\infty} \frac{(1-\alpha)E(\tau)}{N(\tau)\Theta} \left\{ \log\left[\frac{N(\tau)}{N(t)}\right] + \Theta - 1 \right\} e^{-R(\tau)+R(t)}\, d\tau$$

$$\geq \frac{(1-\alpha)E(t)}{N(t)\Theta} \int_t^{\infty} \left\{ \left[\int_t^{\tau} g(v)\, dv \right] + \Theta - 1 \right\} e^{-\rho(\tau-t)}\, d\tau \qquad = \text{if } g(t) > 0$$

Substituting for E(t) using (14) above gives the key equilibrium relation:

$$(16) \quad 1 \geq \frac{(1-\alpha)[L/a_R - g(t)]}{\alpha\Theta} \int_t^{\infty} \left\{ \left[\int_t^{\tau} g(v)\, dv \right] + \Theta - 1 \right\} e^{-\rho(\tau - t)}\, d\tau \quad = \text{if } g(t) > 0$$

Any path of g(t) that satisfies (16) at all times constitutes an intertemporal equilibrium.[16]

Before proceeding to the analysis of the steady state and dynamic paths, it is worthwhile examining (16) in further detail so as to develop a better intuition of the mechanisms at work in the model. As shown by the bracketed term on the right-hand side, a rise in the current rate of invention tends to lower the value of the marginal patent. This effect occurs because, with a reallocation of labour to the inventive sector, the size of the final goods market (given by E) declines. Along any dynamic path consumer expenditure must evolve smoothly according to equation (7), with all changes in consumer expenditure accruing at the cost of exactly counter-balancing changes in R. Consequently, from the point of view of firms at time t, effective market size from time t to infinity is actually given by E(t). The negative effect of a rise in the rate of invention on aggregate consumer expenditure can be termed the market size effect.

The integral term in (16) captures the market creation effect of inventive activity. As discussed earlier above, a rise in the number of existing technologies has both substitution and complementary effects on the current profit flow of existing patent holders, diverting expenditure away from older products and inputs, but, at the same time, creating new markets for existing inputs. Depending upon where in its life cycle an input is, the net effect may be a rise or fall in its current profit flow. Thus, expectations of the timing of future rates of invention should be a critical component in evaluating the value to the marginal firm of an incremental increase in the rate of inventive at any time in the future. However, given the particular utility function chosen

[16] It is easily shown that if (16) is satisfied, so is the consumer's intertemporal budget constraint.

for this simple model, any increase in the rate of invention generates a compensating decrease in the equilibrium rate of interest, which exactly cancels the consumer expenditure dissipation effect. As can be seen from equation (11), however, the remaining two effects are independent of v. Consequently, in general equilibrium, as shown by the integral term on the right-hand side, the benefit of future inventive activity is independent of where in its life-cycle a firm is located.[17] With the CES preferences of section III, this exact cancelation will not occur. Consequently, anticipations of the timing of future rates of invention will become a critical factor in evaluating the value of future inventive activity.

(C) Steady State(s)

One can solve for the steady state(s) of this model by integrating (16) for a constant value of g(t)=g:

$$(17) \quad 1 \geq \frac{(1-\alpha)[L/a_R - g][g + \rho(\Theta - 1)]}{\alpha\Theta\rho^2} = f(g) \quad = \text{if } g > 0$$

The slope of f(g), positive or negative, indicates whether a steady state increase in the rate of innovation increases or decreases the value of the marginal patent and as such indicates whether inventive complementarity or substitutability dominates in the steady state. It is easily shown that f(g) is concave.[18] Innovative activity creates new markets for existing firms, but only at the cost of a decreased overall market size. As the rate of innovation rises, and the overall market

[17] This independence can be seen more clearly by integrating the integral term by parts, to yield:

$$\int_t^{\infty} \{g(\tau)/\rho + \Theta - 1\} e^{-\rho(\tau - t)} d\tau$$

[18] $f''(g) = \frac{-2(1-\alpha)}{\alpha\Theta\rho^2} < 0.$

size shrinks, the marginal value of additional market creation falls, while the cost of a reduction in overall market size, which is increasing in the existing degree of market creation, rises. For complementary effects to be of consequence in the steady state it is necessary that f'(0)>0. Since the sign of f'(0) is given by the sign of $L/a_R - \rho a(\Theta - 1)$, steady state complementary effects will only be present when the size of the total potential market accessed by invention, when divided by the cost of invention (L/a_R), is large relative to the annuity value of the initial market faced by firms in the absence of any further innovation.[19] In the discussion which follows, I shall assume that this condition holds.

Figure I illustrates the different types of equilibria that might arise in the model. For small values of L/aR, the f(g) curve lies everywhere below 1, i.e. regardless of the rate of innovation the profit flow accruing to innovators is insufficient to justify their research costs. In this case, the only steady state is the stagnant equilibrium with no innovation (point A). For sufficiently large values of L/a_R, however, the apex of the f(g) curve attains a tangency with the horizontal line drawn at 1 (point B), which, along with the stagnant steady state (C), constitutes a potential steady state equilibrium. In the absence of further innovation, the initial market facing each research firm is too small to cover the cost of innovation. If, however, there is sustained inventive activity, the market creation effect will generate sufficient profits to allow each successive innovating firm to recoup its research costs. Thus, the economy's equilibrium rate of growth will depend upon whether inventors are pessimistic or optimistic about future rates of

[19] The restriction that Θ be strictly less than 2 in (13) earlier was implied by the partial equilibrium analysis, which did not take into account the general equilibrium interest rate adjustments induced by consumer behavior. More generally, it is simply necessary that Θ not be too large, in the sense identified in the text above. I should note, however, that for intertemporal consumer utility to be bounded it is necessary that $\rho > g$. $\rho > L/a_R$ is sufficient, but not necessary, to guarantee this. If this condition is satisfied, then $\Theta < 2$ is, indeed, necessary for complementary effects to be present in the steady state.

Figure I: Possible Steady State Equilibria

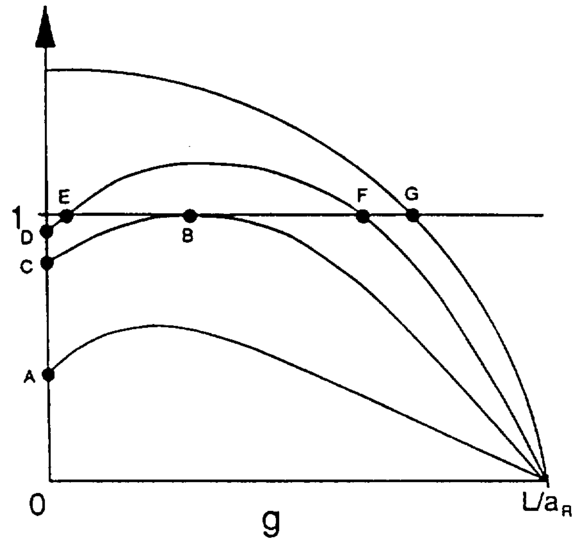
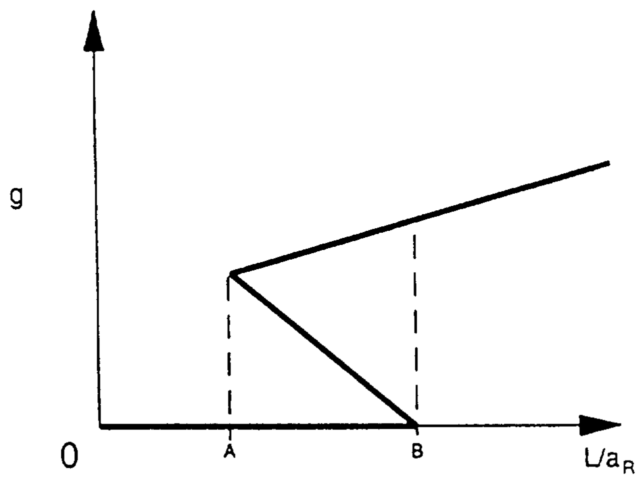


Figure II: Steady State Growth Rates and Relative Market Size



inventive activity. For yet larger values of L/a_R, the f(g) curve cuts the horizontal line twice, and three steady states exist: stagnation (D), a low growth steady state in which inventive activity is locally complementary (E), and a high growth steady state in which inventive activity is locally substitutable (F). Once again, the equilibrium attained by the economy depends upon the expectations of inventors. Finally, for very large values of L/a_R, f(0) is greater than one, and, consequently, f(g) cuts the horizontal line at 1 only once, from above. In this case, where, even in the absence of any further innovation, each research firm can easily recoup the cost of invention, there is only one unique steady state equilibrium (G).

Figure II summarizes the relation between the set of equilibrium steady state growth rates and the relative (to the costs of invention) size of the total potential market. For either very small (<<A) or very large (>B) levels of relative market size, the model behaves exactly like a standard model of invention. In both cases, there exists a unique steady state growth rate which is continuous in the parameters of the model. As will be seen in the next section, under these circumstances any dynamic path which begins at a non-steady state growth rate is easily shown to violate transversality conditions. Thus, an immediate movement to the steady state constitutes the only intertemporal equilibrium. There is no indeterminacy in the model and the economy's growth rate is not dependent upon any coordination of expectations. In both cases, small policy interventions, such as a small subsidy to inventive activity (which is equivalent to moving the horizontal line in figure I down), have small and continuous effects on the economy's equilibrium growth rate.

Figure II also shows, however, that, unlike typical models of invention, for intermediate levels of relative market size this model reveals important discontinuities and indeterminacies. When the market is sufficiently large to allow the f(g) curve to attain a maximum at 1, i.e. at point A in the figure, a positive growth rate, in which the return to each inventor is crucially dependent upon a sustained positive rate of innovation, suddenly appears. At this point, the model begins to resemble a dynamic version of Murphy, Shleiffer and Vishny's (1989) "Big Push" model, in which pecuniary externalities may generate situations in which firms find it profitable to undertake an investment in advanced industrial technologies only when other firms to do so as well.[20] The sudden emergence of this high growth equilibrium also generates an important threshold effect. When the f(g) curve lies slightly below the horizontal line at 1, small policies designed to enlarge the potential market or subsidize the costs of innovation could dramatically transform the economy's opportunity set by allowing the f(g) curve to attain a tangency with the horizontal cost of invention leading, given the appropriate modification of expectations, to a dramatic transition from economic stagnation to rapid growth.

More generally, for values of L/a_R in (A,B) in figure II, the model possesses three steady states, with the selection of an equilibrium crucially dependent upon the coordination of expectations. Interestingly, the two steady states display completely opposite comparative static properties. A small increase in relative market size, or a small subsidy to invention, raises the value of the substitution dominated growth rate, while lowering the value of the complementarity dominanted growth rate. This result is not, perhaps, very surprising. If one raises the return to an endogenous activity in a situation in which the return to the economic actors is locally decreasing in their effort level, then a return to equilibrium will require an increase in their level of activity. If, however, the return to the economic actors is locally increasing in their effort level, as is the case near the complementary steady state, then a return to equilibrium requires a paradoxical reduction in their level of activity.

The contrasting comparative static properties of the different equilibria of this model high-

[20] The "Big Push" model also has the property that for very small or very large market sizes only one, unique equilibrium exists.

light the importance of equilibrium selection. As is well known from the work of Krugman (1991) and Matsuyama (1991), introducing adjustment costs into a model with multiple equilibria does not eliminate the indeterminacy created by the complementarities in the model. More generally, the analysis of the stability of different equilibria, both with complete and incomplete information, has been used to develop opinions as to the equilbrium most likely to prevail in practice.[21] To this end, and so as to gain a deeper understanding of the workings of the model, the following section turns to an analysis of perfect foresight dynamics.

(D) Perfect Foresight Dynamics

In this section I explore the possibility of out of steady state paths for g(t). Since consumer optimality requires that along any perfect foresight path consumer expenditure be smooth and differentiable, it follows, from (14), that any path for g(t) should also be smooth and differentiable. In addition, all such paths must satisfy (16). Consequently, we can simply differentiate (16) with respect to time, which[22] yields:

$$(18) \quad \dot{g} = \rho(L/a_R - g)\left\{1 - \frac{(1-\alpha)[L/a_R - g][g + \rho(\Theta - 1)]}{\alpha\Theta\rho^2}\right\}$$

Or:

$$(18)' \quad \dot{g} = \rho(L/a_R - g)[1 - f(g)]$$

It is easily shown that any path for g(t) which satisfies (18)' and which converges to a positive steady state will satisfy (16) at all points in time and, consequently, be an intertemporal

[21] See, for examples, Blanchard & Fischer (1989, ch. 5), Grandmont (1985) and Lucas (1986).

[22] Substituting where necessary using the fact that (16) must hold with equality along any dynamic (i.e. not identically equal to zero) path for g(t).

equilibrium.[23]

Figure III below explores the dynamics of g(t) for two particular cases of the model. In (a) we see that when only one, substitution dominated, steady state exists all out of steady state dynamic paths diverge to either 0 or L/a_R. Both types of paths ultimately violate the free entry condition, and hence cannot constitute perfect foresight equilibria. The divergence of these out of steady state paths in standard, substitution dominated, models is typically used to justify the immediate movement to the steady state,[24] which all actors know is the only potential intertemporal equilibrium. In (b), however, we see that when an additional complementary steady state exists the divergent paths emanating from the substitution dominated steady state actually converge to the complementary steady state, which is stable. In this case, any dynamic path with an initial value of g less than or equal to g_s constitutes a legitimate perfect foresight equilibria. Only paths beginning above g_s, which are divergent, can be ruled out.

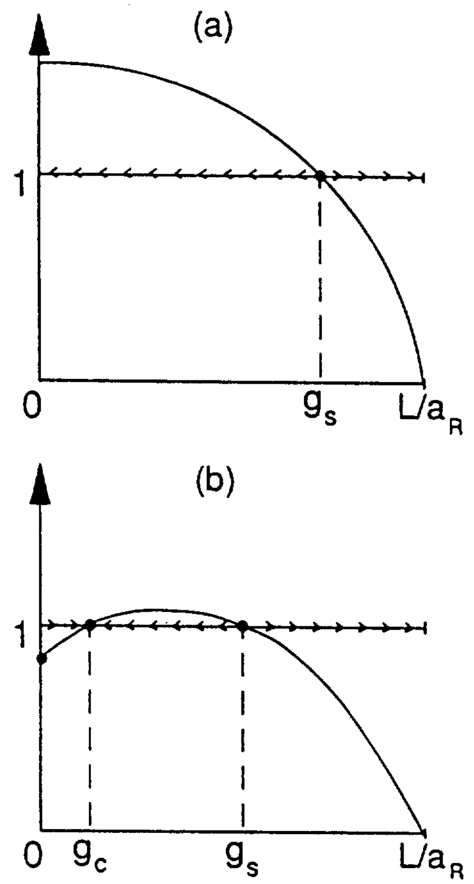
The intuition behind the dynamics which underly these perfect foresight paths is best understood by examining the equilibrium behavior of the interest rate, $\dot{R}$. If V(x,t) is the market value of the patent to good x at time t, then, under perfect foresight, it must be the case that:

$$(19) \quad \dot{R}(t) = \frac{\pi(x,t)}{V(x,t)} + \frac{\dot{V}(x,t)}{V(x,t)}$$

[23] Intuitively, since (18)' is derived by differentiating (16), any path given by (18)' satisfies (16) up to a constant of integration. If the path converges to a positive steady state, then, in the limit, g is constant and, by the definition of the steady state, must satisfy (16). Consequently, the implied constant of integration is zero.

[24] See Grossman and Helpman (1991b, chs. 3 and 4).

Figure III: Dynamic Paths for g(t)



i.e. the (common) risk free rate of return on each asset equals profit flows plus patent value appreciation divided by the current patent value. Applying this condition to V(N(t),t), the marginal patent, yields:[25]

$$(19)' \quad \dot{R} = \frac{(1-\alpha)(L/a_R - g)(\Theta - 1)}{\alpha\Theta} - g - \frac{(1-\alpha)(L/a_R - g)g(\Theta - 1)}{\Theta\alpha\rho} + \frac{(1-\alpha)(L/a_R - g)g}{\alpha\rho}$$

$$= \frac{(1-\alpha)(L/a_R - g)(\Theta - 1)}{\alpha\Theta} - g - \frac{(1-\alpha)(L/a_R - g)g}{\Theta\alpha\rho}$$

The first term on the right hand side is the flow of current profits divided by the value of the marginal patent and is clearly decreasing in g. A rise in the rate of invention implies a smaller final goods market size, and, by extension, a smaller flow of current profits. The next three terms equal the proportional rate of patent value appreciation. As emphasized earlier above, the invention of new inputs dissipates consumer expenditure across a broader range of products and, consequently, lowers the value of the marginal patent. This is reflected in the -g term.[26] The creation of new inputs, also leads to a substitution by current producers away from existing inputs. This effect is reflected in the third term. Finally, the invention of new inputs creates new markets for input N(t), as reflected in the fourth term. The flow value of both these last two effects is related to the product of the existing aggregate market size (L/a_R-g) times the rate of market creation/producer substitution (g) and, hence, is non-monotonic in g. Division by ρ reflects the net present value of these flow values. The second line of the expression combines the last two effects, highlighting the fact that the positive effect of market creation effect always

[25] (19)' can be derived by differentiating (15) with respect to time.

[26] The compensating interest rate adjustment discussed earlier applies to the effect on the net present value of future profits of an increase in current and future rates of inventive activity. An increase in the number of inventions actually available at any point in time does, however, lower the value of all patents. Examine, for example, equation (15)'.

dominates the loss due to producer substitution.[27]

Along any perfect foresight path, $\hat{E} = \dot{R} - \rho + g$. Focusing on case (b) in figure III, when g is less than g_c the current profit flow is large, but the market creation element of patent value appreciation is small. Consequently, $\dot{R} + g$ is less than ρ and expenditure falls, leading to a higher rate of invention and a convergence to g_c. When g is in (g_c, g_s), the market creation effect is reasonably large, as is the current profit flow. Consequently, $\dot{R} + g > \rho$, and expenditure rises, leading to a reduction in the rate of invention and a convergence to g_c. Finally, when g is large and greater than g_s, market size is small and, consequently, both the current profit flow and the net present value of market creation are small. Thus, $\dot{R} + g < \rho$, leading to a fall in expenditure, and a divergence away from the substitution dominated steady state. In this manner, the interest rates implied by the current profit flow and patent value appreciation determine the stability and instability of the different steady states.

The stability of the complementary steady state should not, necessarily, lead one to conclude that it is the most likely long run equilibrium. Recent work on using learning as a means of choosing amongst multiple equilibria in overlapping generations models, e.g. Grandmont 1985 and Lucas 1986, has shown that those equilibria usually deemed to be unstable (stable) under certainty are frequently stable (unstable) in a learning environment, where agents use past information to forecast the future behavior of the system.[28] Admittedly, as argued by Lucas, there is no reason to believe that these results are independent of the ad hoc structure of the learning process selected to close these models. Nevertheless, one might argue that under more realistic informational assumptions the substitution dominated steady state would actually be sta-

[27] See footnote 15 earlier above for an explanation.

[28] In essence, the dynamic equations governing consumer and producer behavior in the economy are reversed, which reverses the stability of all the equilibria.

ble, and hence the most likely candidate for a long run equilibrium. Rather than examine the dynamics of the model in a learning environment, I prefer to remind the reader that a highly particular set of preferences were used to simplify and close this particular version of the model. As will be seen in the following section, the use of fairly standard CES preferences enormously expands the range of dynamics possible including, for appropriate parameter values, circumstances in which *both* the substitution and complementary dominated steady state are stable. In this case, neither perfect foresight nor learning dynamics provide much indication as to which equilibrium would be selected.

III ALLOWING FOR LIFE CYCLE EFFECTS IN GENERAL EQUILIBRIUM

Much of the analytical simplicity of the preceeding pages stems from the unusual preferences selected to close the model, where an unusually strong preference for variety conspired, in general equilibrium, to cancel many of the life cycle effects inherent to the model's production structure. This section shows how the use of standard CES preferences brings to the fore all of the life cycle effects, which, while not changing the number and type of steady states, enormously expands the potential set of out of steady state dynamic paths.

Let instantaneous utility now be given by the usual CES formulation:

$$(20)\quad V(t)=\left[\int_0^{N(t)} C(s)^{\eta}\,ds\right]^{\frac{1}{\eta}}\qquad \eta>0$$

where C(s) denotes consumption of final good s. Let total intertemporal utility be given by:

$$(21)\quad U=\int_t^{\infty} \ln[V(\tau)]e^{-\rho(\tau-t)}\,d\tau$$

I shall keep the production side of the economy largely unchanged, except that I now assume that it takes one unit of labour to make one unit of any input v at any time. In order to simplify the algebra I shall also assume that Θ equals one. With this assumption, each research firm depends entirely on further innovation to create a market for its inventions. Consequently, complementarity will always be dominant at low growth rates and, for a sufficiently large total market size, multiple equilibria will always arise.

(A) General Equilibrium

Since much of the general equilibrium structure parallels that of the simple model in section II, I shall keep the presentation brief. As in the simple model, the demand for input v by producers of final good s is given by:

$$(22)\quad x^D(v,s) = \frac{p(v)^{-\epsilon}E(s)}{\int_0^s p(v)^{1-\epsilon}\,dv} \qquad \epsilon = \frac{1}{1-\alpha}$$

With one unit of labour required to produce one unit of any input v, intermediate input producers will find it optimal to charge a price:

$$(23)\quad p(v) = \alpha^{-1}\ \forall\, v$$

where I have set the nominal wage equal to one. Consequently, in equilibrium:

$$(24)\quad x^D(v,s) = \bar{x}(s) = \alpha E(s)/s$$

The equilibrium (perfectly competitive) price of final good s is then easily computed to be:[29]

$$(25)\quad p(s) = \frac{s^{\frac{\alpha-1}{\alpha}}}{\alpha}$$

Given their CES preferences, total consumer demand for each final good s will be given by:

$$(26)\quad Q^D(s) = \frac{p(s)^{-\tau}E}{\int_0^N p(s)^{1-\tau}} \qquad \tau = \frac{1}{1-\eta}$$

where E is aggregate instantaneous expenditure. With equilibrium final good prices given by (25) above, it follows that total consumer expenditure on each final good s is given by:

$$(27)\quad E(s) = \frac{(\psi+1)s^{\psi}E}{N^{\psi+1}} \qquad \psi = \frac{(1-\alpha)\eta}{\alpha(1-\eta)} > 0$$

[29] With zero profits, p(s)Q(s)=E(s). (25) then follows from the fact that:

$$Q(s) = \left[\int_0^s x(v,s)^{\alpha}\right]^{\frac{1}{\alpha}} = \left[\int_0^s \bar{x}(s)^{\alpha}\right]^{\frac{1}{\alpha}} = s^{\frac{1}{\alpha}}\bar{x}(s) = s^{\frac{1-\alpha}{\alpha}}\alpha E(s)$$

As in the model of section II, a rise in the available number of final goods, N, will lead to a dissipation of consumer expenditure expenditure across a broader range of products and, consequently, a reduction in total expenditure per final good. With logarithmic intertemporal utility, however, the time path of optimal expenditure satisfies:

$$(28) \quad \hat{E}(t) = \dot{R}(t) - \rho$$

Thus, unlike the previous model, in general equilibrium a rise in the rate of invention will not generate a compensating decrease in the interest rate.

Using (24) and (27), the instantaneous profits of each intermediate input producer v are easily found to be:

$$(29) \quad \pi(v) = \frac{(1-\alpha)(\psi+1)E}{\psi}\left\{\frac{N^{\psi} - v^{\psi}}{N^{\psi+1}}\right\}$$

Taking the partial (equilibrium) derivative of this expression with respect to N:

$$(30) \quad \frac{\partial \pi(v)}{\partial N} = -\frac{(\psi+1)\pi(v)}{N} + \frac{(1-\alpha)(\psi+1)E}{N^2}$$

As in the model of section II, the first (negative) term represents the loss due to the dissipation of consumer expenditure across a broader set of final goods, while the second (positive) term represents the gain due to the creation of additional markets for each input.[30] Once again, each input will pass through a life-cycle, with the market creation effect dominating early in an invention's life and the expenditure dissipation effect dominating as the technology matures:

[30] Since $\Theta = 1$ existing producers cannot make use of new inputs and, consequently, there is no longer a term reflecting the loss due to substitution by existing producers towards the new technologies.

$$(31)\quad \frac{\partial \pi(v)}{\partial N} > 0 \quad \forall\, v > N(\psi+1)^{-\frac{1}{\psi}}$$

$$\frac{\partial \pi(v)}{\partial N} < 0 \quad \forall\, v < N(\psi+1)^{-\frac{1}{\psi}}$$

Finally, as before, free entry into invention requires that the value of the marginal patent be less than or equal to the cost of invention:

$$(32)\quad \frac{a_R}{N(t)} \geq \int_t^{\infty} \pi(N(t),\tau)e^{-R(\tau)+R(t)}\, d\tau$$

Or, substituting and simplifying:

$$(32)'\quad 1 \geq \frac{(1-\alpha)(\psi+1)[L/a_R - g(t)]}{\alpha\psi} \int_t^{\infty} \left\{ e^{-\int_t^{\tau} g(v)\,dv} - e^{-(\psi+1)\int_t^{\tau} g(v)\,dv} \right\} e^{-\rho(\tau-t)}\, d\tau \quad = \text{if } g(t) > 0$$

Where I have used the fact that labour market clearing requires, as in the previous model, that $L = \alpha E + g a_R$. Any path for g(t) that satisfies (32)' at all times is a perfect foresight intertemporal equilibrium.

It is worth examining (32)' in detail to gain a better understanding of the mechanisms at work in this, more complicated, version of the model. As in the simple model of section II, a rise in the current rate of invention implies a reduction in the total labour in manufacturing, and, consequently, a reduction in current (and, by continuity, future) consumer expenditure, which lowers the value of the marginal patent. This is reflected in the bracketed term in front of the integral. Unlike the simple model of section II, however, there is no interest rate adjustment compensating for the reduction in consumer expenditure per product brought about by a rise in the number of available technologies. Thus, the life cycle tension of complementarity and substitution remains prominent in the model, as can be seen from the integral term in (32)'. As long as

$e^{\int_t^\tau g(v)\,dv} < (\psi+1)^{1/\psi}$ the term in curlicues is increasing in the cumulative integral of g, i.e. in the cumulative increase in the number of products. Once this critical value is passed, however, the term in curlicues is decreasing in the cumulative integral of g.[31] This reflects the life cycle transition from complementarity to substitution, highlighted in (31) earlier above.

(B) Steady State(s)

One solves for the steady state by integrating (32)' for a constant value of g, which yields:

$$(33) \quad 1 \geq \frac{(1-\alpha)(\psi+1)(L/a_R - g)g}{\alpha[\rho+g][\rho+g(\psi+1)]} = f(g) \quad = \text{if } g > 0$$

As before, f(g) is a concave function. Given that I have set the parameter Θ equal to 1, f(0)=0, i.e. the marginal inventor depends crucially upon further invention to reap a positive return. Consequently, complementary effects will always be present and for a sufficiently large L/a_R the model will have three steady states: stagnation with g=0, a positive low growth steady state in which inventive complementarity dominates (on average) over the product life cycle, and a positive high growth steady state in which inventive substitutability dominates on average.

We can solve algebraically for the steady state growth rates by rearranging (33) so that:

$$(34) \quad ag^2 + bg + c = 0$$

where: $a = (\psi+1)/\alpha$

$$b = \rho(\psi+2) - \left(\frac{1-\alpha}{\alpha}\right)(\psi+1)(L/a_R)$$

$$c = \rho^2$$

[31] Ideally, the inventor of product N(t) would like invention to proceed at a rapid pace until a maximal market size is obtained, when $N(t)/N(\tau) = (\psi+1)^{-1/\psi}$, at which time he would like all further invention to cease.

Clearly, for this equation to have two positive real solutions it is necessary and sufficient that $b<0$ and $b^2-4ac>0$. This information will prove useful in our analysis of expectational dynamics in the next section.

(C) Perfect Foresight Dynamics

As before, we can analyze out of steady state paths by differentiating (32)' with respect to time. Doing so[32] provides the following equation governing the dynamics of g(t):

$$(35)\quad \dot{g}(t) = [L/a_R - g(t)]\left\{\rho + g(t)(\psi+1) - \left(\frac{1-\alpha}{\alpha}\right)(\psi+1)[L/a_R - g(t)]g(t)F(t)\right\}$$

where:

$$(36)\quad F(t) = \int_t^{\infty} e^{-\int_t^{\tau} g(v)\,dv} e^{-\rho(\tau-t)}\,d\tau$$

(35), with its dependence on the integral of future values of g, is a non-linear integro-differential equation and, hence, cannot be analyzed with a one dimensional phase diagram. Further differentiation of (35), however, can reduce the problem to one of analyzing a second order differential equation in g using a two-dimensional phase diagram. Analytically, however, it is considerably simpler to note that:

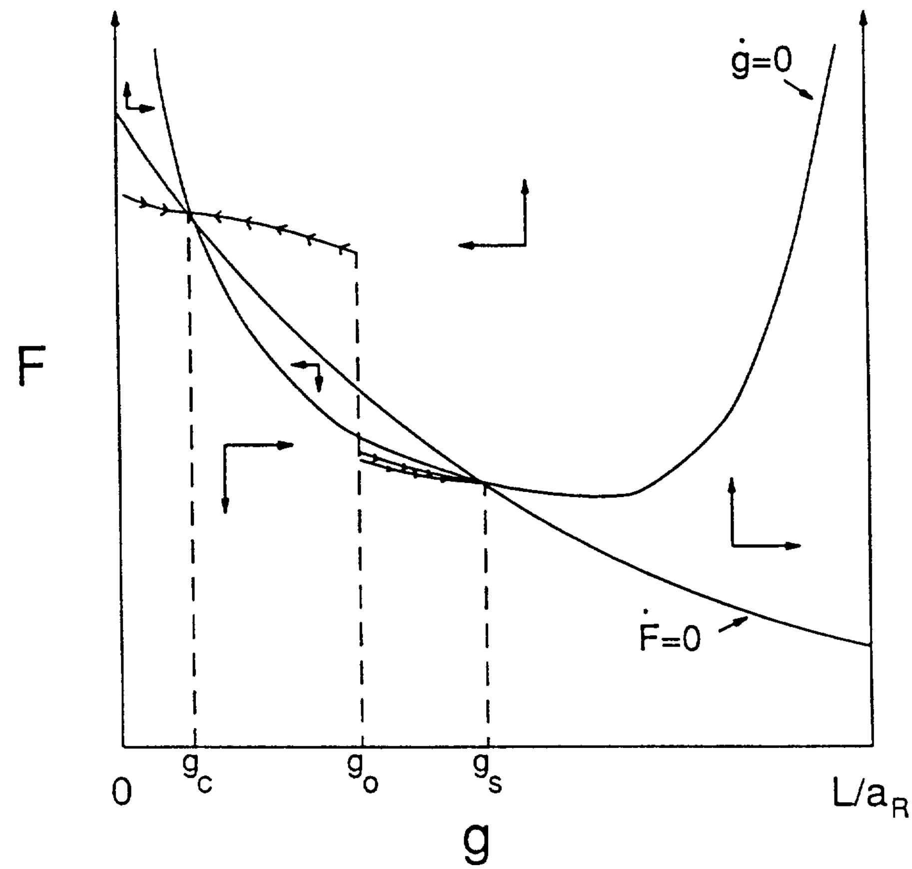
$$(37)\quad \dot{F}(t) = [\rho + g(t)]F(t) - 1$$

The analysis of the dynamics of F and g in the two dimensional plane is equivalent to the analysis of the second order differential equation in g.

Figure IV presents the phase diagram for the F,g system. From (37) it is not hard to see that the $\dot{F}=0$ locus is downward sloping, with $\dot{F}$ greater than (less than) zero everywhere above (below) the locus. In contrast, the $\dot{g}=0$ locus is convex and asymptotes to the F axis and a verti-

[32] Substituting where necessary using the fact that (32)' must hold with equality along any positive growth path.

Figure IV: Dynamics of F and g



cal line at g=L/a_R. Since $\partial \dot{g}/\partial F$ is clearly less than zero, we know that everywhere above (below) this locus $\dot{g}$ is less than (greater than)-zero. For the sake of brevity, I focus on the case where there are two positive growth steady states, i.e. two stationary points.

We begin our analysis by linearizing the system around the steady state. Let:

$$\mathrm{A} = \begin{pmatrix} \partial \dot{F}/\partial F & \partial \dot{F}/\partial g \\ \partial \dot{g}/\partial F & \partial \dot{g}/\partial g \end{pmatrix}$$

denote the matrix of partials of the dynamic system. With some algebra, it can be shown that the determinant of this system evaluated at the steady states equals:

$$(38) \quad |\mathrm{A}|_{SS} = (L/a_R - g_{SS})\left[\ \pm\sqrt{b^2 - 4ac}\ \right]$$

where a, b and c are the coefficients of the quadratic equation (34) used to determine the steady state growth rates earlier above. Since the negative root of $\sqrt{b^2 - 4ac}$ corresponds to the slower (complementary) steady state growth rate, we immediately know that the complementary steady state is saddlepath stable, as drawn in figure IV. The stability of the substitution dominated steady state will depend upon the trace of the system, which equals:

$$(39) \quad tr(\mathrm{A}_{SS}) \ = \ g(\psi + 2) - \frac{L\rho}{a_R g} + 3\rho$$

Appendix I shows that this trace may be either negative or positive. Thus, locally the substitution dominated steady state may be either stable or unstable.[33]

Clearly, with the incorporation of preferences which do not cancel the life cycle effects inherent to its the production structure, this model allows for a far greater variety of expectationally induced paths. Even if one takes as given an initial value of g, the subsequent behavior of the system can still be largely indeterminate, since a variety of initial values for F can be

---

[33] In addition, g_s can be either a node, a focus or a centre.

self-justified by the subsequent dynamics. For example, consider an initial value of g (g_o) in the neighbourhood of the substitution dominated steady state, which I will assume is stable in this case.[34] If the economy selects a large value of F, on the saddlepath to the complementary dominated steady state, then the rate of invention will gradually fall and the the economy will converge to g_c. If, however, the economy selects a small value of F, then the rate of invention will rise and the economy will converge to the substitution dominated steady state, g_s. Different expectations of future rates of invention, which determine F, can now give rise to different self-fulfilling equilibrium paths.

The dynamic behavior of this model is, once again, most easily understood by analyzing the determinants of the equilibrium interest rate. As before, under certainty the equilibrium common rate of return on all assets equals profit flows plus patent value appreciation. With Θ equal to 1, input N(t) earns zero profits. Consequently, the equilibrium rate of interest is simply given by the instantaneous rate of appreciation of the patent to input N(t):

$$(40) \quad \dot{R} = -g(\psi+1) + (\psi+1)(1-\alpha/\alpha)(L/a_R - g)gF$$

The first term on the right-hand side represents the loss in patent value due to the dissipation of consumer expenditure across a broader range of products brought about by the invention of new intermediate inputs. The second term on the right hand side is the net present value of market creation, which is now seen to depend upon the market size (L/a_R-g), the rate of market creation (g), and the discount factor F. The discount factor is no longer simply the inverse of the rate of time preference, but rather a more complicated term which takes into account the rate and timing of future innovation.

[34] For g_s to be stable, it is necessary that the $\dot{g} = 0$ locus be downward sloping at the point of intersection, as drawn in the figure. I have drawn g_s as being a stable node, although it could just as easily be a stable focus.

If research firms anticipate that invention, although high now, will fall rapidly in the near future, then the net present value of current market creation is large (as F is large), as firms anticipate that for some time they will be in a position where profits are strictly increasing in the number of inventions. This leads to a high value of patent appreciation, which makes $\dot{R} > \rho$, leading to a rapidly rising path of expenditure and a rapidly falling rate of invention. This is illustrated by the saddlepath convergence from g_o to g_c in figure IV. If, however, the marginal firm anticipates rapid rates of invention in both the near and long term future, then the net present value of current market creation is small, as the firm anticipates that it will soon transit to maturity, where profits are strictly decreasing in the total number of inputs. Consequently, patent value appreciation is small, which makes $\dot{R} < \rho$ and leads to a falling path of expenditure and a rising rate of invention. This is illustrated by the many paths converging from g_o to g_s in figure IV, any of which constitutes an intertemporal equilibrium.

This section has shown how preferences which do not cancel the life cycle effects inherent to the model greatly expand the range of indeterminacy, allowing a variety of expectations of future inventive activity to constitute legitimate self-fulfilling intertemporal equilibria. When both equilibria are stable, perfect foresight dynamics provide little indication as to which equilibrium will be selected. Neither, for that matter, is the conventional wisdom of learning dynamics, with its reversal of stability, helpful, since it implies there will be convergence to neither equilibrium. One might conjecture, however, that in this case there exists an unstable limit cycle, which would be stable under learning.[35] Lest the reader draw solace from this fact, and its implication that one should not worry about the perverse comparative statics of the complementary steady state, I should once again emphasize that the dynamics of the model are extraordi-

[35] I have not been able to prove such a limit cycle exists.

narily sensitive to the selection of preferences. For example, with instantaneous CES preferences with an infinite intertemporal elasticity of substition, the dynamics are exactly the reverse of those shown in figure III earlier above: the complementary equilibrium becomes unstable, while the substitution dominated equilibrium becomes stable.

IV CONCLUDING REMARKS

Like many new technologies, this paper is both creative and destructive. Inspired by historical examples of a life cycle transition from complementarity to substitution, I have presented a simple model in which new technologies destroy the rents accruing to older technologies by inducing a substitution of producer and consumer expenditure towards new products and inputs, while, at the same time, creating rents for older technologies by introducing new applications and markets for these older inputs. The tension between these two effects produces a stylized life cycle, in which new technologies initially find further invention complementary, but then, as they develop mature and large markets, find that the dominant effect of further invention is substitution. The model generates an interesting range of equilibria, including threshold effects where small policies might have large consequences and conditions under which there exist multiple steady states, including one, the complementarity dominated steady state, in which the effects of policies are decidedly perverse.

With the exception of the linear models of factor accumulation,[36] almost all contemporary models of endogenous growth are built around external effects. There is no need to assume, however, as has been the practice in existing models, that the individual return is monotonically related to the aggregate level of activity. Allowing for non-monotonic external effects immediately creates the possibility of multiple equilibria, with all of the indeterminacies they entail. Since each of these equilibria display different comparative static properties, the implications of policy actions are dependent upon which equilibrium is selected. Equilibrium selection is, however, a non-trivial problem. Adjustment costs are of little help in narrowing down the potential set, while the implications drawn from an analysis of dynamics (either with full or incomplete

[36] E.g. Rebelo (1991) and Jones and Manuelli (1990).

information) depend upon a certain stability in the behavior of the model. As shown in the analysis of perfect foresight dynamics in sections II and III above, the dynamic behavior of the model presented in this paper is extraordinarily dependent upon the preferences chosen to close the model. The range and type of dynamics is so great that it is hard to imagine that any analysis of dynamics, either with or without complete information, will help select an equilibrium. This paper has, by no means, provided any answers to the problem of equilibrium selection. If, however, models of endogenous growth are to be built around external effects, it is an issue which they must surely, sooner or later, confront.

V BIBLIOGRAPHY

Aghion, Philippe and Howitt, Peter. "A Model of Growth through Creative Destruction." Econometria 60, No. 2 (March 1992): 323-351.

Blanchard, Olivier J. and Fischer, Stanley. Lectures on Macroeconomics. Cambridge, MA: The MIT Press, 1989.

Burke, James. The Day the Universe Changed. Boston: Little, Brown and Company, 1985.

Graham, Gerald S. "The Ascendancy of the Sailing Ship 1850-85." Economic History Review, 2nd series, 9 (August 1956): 522-530.

Grandmont, Jean-Michel. "On Endogenous Competitive Business Cycles." Econometrica 53 (September 1985): 995-1045.

Grossman, Gene M. and Helpman, Elhanan. "Quality Ladders in the Theory of Growth." Review of Economic Studies 58 (1991a): 43-61.

Grossman, Gene M. and Helpman, Elhanan. Innovation and Growth in the Global Economy (Cambridge, MA: MIT Press, 1991b).

Harley, Charles K. "The shift from sailing ships to steamships, 1850-1890: a study in technological change and its diffustion." In Donald N. McCloskey, ed. Essays on a Mature Economy: Britain after 1840. Princeton, N.J.: Princeton University Press, 1971.

Haites, Erik F. and Mak, James. "Ohio and Mississippi River Transportation 1810-1860." Explorations in Economic History 8, No. 2 (Winter 1970/71): 153-180.

Jones, Larry E. and Manuelli, Rodolfo. "A Convex Model of Equilibrium Growth." Journal of Political Economy 98, No. 5 (1990): 1008-1038.

Krugman, Paul. "History versus Expectations." Quarterly Journal of Economics 106, No. 2 (May 1991): 651-667.

Law, R.J. The Steam Engine: A brief history of the reciprocating engine. London: HMSO Publications, 1965.

Lucas, R.E. Jr. "Adaptive Behavior and Economic Theory." Journal of Business 59 (1986): S401-S426.

Mak, James and Walton, Gary M. "Steamboats and the Great Productivity Surge in River Transportation." Journal of Economic History XXXII (1972): 619-640.

Matsuyama, Kiminori. "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium." Quarterly Journal of Economics 106, No. 2 (May 1991): 617-650.

Murphy, Kevin M.; Shleifer, Andrei; and Vishny, Robert. "Industrialization and the Big Push." Journal of Political Economy 97 (1989): 1003-26.

Rebelo, Sergio. "Long Run Policy Analysis and Long Run Growth." Journal of Political Economy 99 (June 1991): 500-521.

Romer, Paul M. "Endogenous Technological Change." Journal of Political Economy 98, No. 5, p. 2 (October 1990): S71-S102.

Romer, Paul M. "Growth Based on Increasing Returns Due to Specialization." American Economic Review 77 (May 1987, Papers & Proceedings): 56-62.

Rosenberg, Nathan. Perspectives on Technology. Cambridge: Cambridge University Press, 1976.

Young, Alwyn. "Invention and Bounded Learning by Doing." Manuscript, 1992. Forthcoming. Journal of Political Economy.

VI APPENDIX I: STABILITY OF THE SUBSTITUTION-DOMINATED STEADY STATE EQUILIBRIUM

As noted in the text, the stability of the substitution dominated equilibrium in the general model depends upon the sign of the trace of the matrix of partials of the F,g dynamic system evaluated at the steady state:

$$(A.1) \quad tr(A_{SS}) = g(\psi+2) - \frac{L\rho}{a_R g} + 3\rho$$

This appendix will show that this trace may be either positive or negative at either of the positive growth steady states.

I begin by noting that the two steady state growth rates and (A.1) are continuous functions of all of the parameters of the model. Furthermore, I remind the reader that $b>0$ and $b^2-4ac>0$ (see (33) in the text) are necessary for the model to have more than one steady state. As b^2-4ac goes to zero (from above), the complementary and substitution steady state growth rates converge to a common value, that of a unique positive growth steady state. Thus, if I can show that in the case of that unique positive growth steady state (A.1) is strictly positive (strictly negative) for a given set of parameter values, it follows that in a neighbourhood of those parameter values such that $b^2-4ac>0$, i.e. such that two positive growth equilibria exist, (A.1) will still be strictly positive (strictly negative).

A unique positive growth equilibrium only exists when $b^2-4ac=0$. Expanding this expression and solving for L/a_R as a function of the other parameters:

$$(A.2) \quad \frac{L}{a_R} = \rho\alpha\left\{\frac{\psi+2\pm 2\sqrt{\frac{\psi+1}{\alpha}}}{(1-\alpha)(\psi+1)}\right\}$$

Since $b>0$ is necessary for a positive growth equilibrium, we take the positive root:

$$(A.2)' \quad \frac{L}{a_R}|_{SS_{unique}} = \rho\alpha\left\{\frac{\psi + 2 + 2\sqrt{\frac{\psi+1}{\alpha}}}{(1-\alpha)(\psi+1)}\right\}$$

Since g=b/2a, we have:

$$(A.3) \quad g_{SS_{unique}} = \rho\sqrt{\frac{\alpha}{\psi+1}}$$

Substituting (A.2)' and (A.3) into (A.1):

$$(A.4) \quad tr(A_{SS_{unique}}) = \rho\left\{(\psi+2)\left(\sqrt{\frac{\alpha}{\psi+1}}\right)\left\{1 - \frac{1}{1-\alpha}\right\} + 3 - \frac{2}{1-\alpha}\right\} = g(\alpha)$$

Clearly, g(0)>0, while $g(\alpha) < 0$ for all $\alpha \geq 1/3$. Thus, the trace of A may be strictly positive or strictly negative in the case of the unique positive growth steady state. By continuity, it follows that the trace of A may be strictly positive or strictly negative at either of the two multiple equilibrium positive growth steady states.