NBER WORKING PAPERS SERIES

R&D, INVESTMENT AND INDUSTRY DYNAMICS

Saul Lach

Rafael Rob

Working Paper No. 4060

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1992

This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #4060
April 1992

R&D, INVESTMENT AND INDUSTRY DYNAMICS

ABSTRACT

We present a model of industry evolution where the dynamics are driven by a process of endogenous innovations, followed by subsequent embodiments in physical capital. Traditionally, the only distinction between R&D and physical investment was one of labeling: the first process accumulates an intangible stock (knowledge) while the second accumulates physical capital; both stocks affect output in a symmetric fashion. We argue that the story is not that simple, and there is more to it than differences in the object of accumulation. Our model stresses the causal relationship between past R&D expenditures and current investments in machinery and equipment. This causality pattern, which is supported by the data, also explains the observed higher volatility of physical investment (relative to R&D expenditures).

Saul Lach
The Hebrew University
Department of Economics
Jerusalem, 91905
Israel
and NBER

Rafael Rob
University of Pennsylvania
Department of Economics
3718 Locust Walk
Philadelphia, PA 19104-6297

1. Introduction

Traditionally, economists distinguish between increases in knowledge and increases in physical capital. While such distinction is useful in classifying the determinants of economic growth (either at the firm level or at the level of the whole economy), the thesis of this paper is that their interdependence can fruitfully explain the life-cycle of an industry, i.e., the sequence of events that occurs between the introduction of a new product and the attainment of a steady-state. For instance, it can explain the pattern of prices, quantities, the number of operating firms, the rate of entry and exit, the rate of patenting, etc. Of particular interest to us here is the relationship between the time series of R&D expenditures and the time series of physical investments. That is, the manner in which the two series co-vary over time, and the determinants of this co-variation.

Empirical investigations of this aspect of industry dynamics exhibit two salient features. First, it has been shown that past R&D expenditures cause (in a statistical sense) current investment expenditures. Second, it has been shown that the volatility of the physical investment series is much higher than its' R&D counterpart. These findings are mostly derived from firm-level data. See Ben-Zion (1984), Mairesse and Siu (1984), Gordon, Shankerman and Spady (1986) , Lach and Shankerman (1989), and Hall and Hayashi (1989). Given such findings our aim here is twofold. First, we empirically confirm that the same features persist at the industry-level as well. Second, we argue that existing theoretical formulations are not capable of rationalizing such features of industry dynamics. In their place, we set up a new model where R&D and physical investments are viewed as vertically linked (see below for details), and where the above evolutionary patterns arise quite naturally.

A brief glimpse at the literature might be useful at this point. A variety of papers use patent-race ideas to explain the persistence of monopoly (Gilbert-Newbery (1982), Reinganum (1983), Fudenberg-Tirole (1986)), the market structure as determined by the R&D process (Reinganum (1985), Flaherty (1980)) , the incentives to invest in R&D as determined by one's position in a Patent race (Harris and Vickers (1985)), the incentive to invest in R&D as determined by the market structure (Kamien and Schwartz (1980), Loury (1979)), the discrepancy between the private and the social incentives to engage in R&D (Loury (1979)), and many more. All these formulations, however, are purely R&D-theoretic. Hence, by their very construction, they cannot explain the relationship between R&D expenditures and physical investments.

At the other end of the spectrum, there is the more measurement-oriented productivity literature. See the volume edited by Griliches (1984). This literature treats knowledge and capital as symmetrical inputs into a common Neoclassical production function. Hence, it can relate properties of this production function (eg. substitutability, complementarity, etc.) to the accumulation paths of R&D and physical investments. This literature, however, encounters difficulties in rationalizing a relationship of causality running from R&D expenditures to physical investments. As mentioned, the formulation views knowledge and capital as horizontally related, i.e., as "simultaneous" inputs into a single production function. Therefore, the two are contemporaneously determined, not sequentially; so there is no relationship of causality. A second difficulty relates to the (empirical) fact that the R&D series is smoother than the investment series. Proponents of the productivity literature are driven to rationalize this by differential costs of adjustments; namely, it must be the case that R&D expenditures are harder to vary than physical investments. It is less than clear, however (on a priori grounds at least), whether that is indeed the case.[1] A third difficulty with the productivity literature is that the formulation is at the individual-firm level; or, equally restrictively, that strong aggregation properties are satisfied. Hence, this formulation is not amenable to industry-level analysis.

Given these difficulties, we suggest a different way of looking at the relationship between R&D and investments. As in the patent-race literature we view R&D as leading (stochastically) to innovations. The latter, however, come in the form of ideas blueprints, prototypes, etc. So- by themselves-they are not directly operational. They need to be implemented. That is, new machinery and equipment which embodies the new ideas must be created and put into place. Only then can one generate a flow of profit. So the basic idea of the formulation is that investments respond to innovations,and that the latter occur randomly. We shall refer to this idea as the implementation (or embodiment) hypothesis.

We embed the implementation hypothesis into an industry context, generating an equilibrium time-path of entry and exit which settles eventually into a steady-state. The nature of the industry's turnover is such that new entrants introduce technologies superior to incumbents' technologies (in terms of the unit-costs of production), thereby replacing them. Each replacement is associated with investments in physical capital. Hence, the "blips" in the investment process. On the other hand, R&D expenditures must be maintained at

[1]Variations in R&D intensity occur via the hiring of more R&D personnel (i.e., scientists and engineers) and, perhaps, through the purchase of additional laboratory equipment. On the other hand, variations in physical investments occur via the purchase and installation of new machines, and through the training of workers to operate them. It is not clear why the marginal cost of the former should rise more steeply than the latter.

a steady rate as long as one hopes to generate (eventually) an innovation. Hence, the relative constancy of the R&D process. So the essence of the formulation is to create a vertical link between the R&D process and the implementation-of-innovations process (ie, investments), and to generate implications concerning their co-evolution. As a by-product of this we also generate implications concerning the general dynamics of the industry.

In regard to the latter the primary results we report are as follows. First, entry and exit occur as new innovations arise, and at certain stages they occur simultaneously. Second, the survivors of this selection process make (possibly) substantial profits. Third, the R&D process is less volatile than the investment process. Fourth, prices decline and quantities increase along the industry's evolutionary path. Most of these implications match the stylized facts presented in Gort and Klepper (1982), Klepper and Graddy (1990), Mueller (1986), Ericson and Pakes (1989), and Jovanovic and MacDonald (1990). Furthermore, as regard the relationship between R&D and investments we present further empirical evidence to support the model's implications.

The paper is organized as follows. The following section presents basic statistics on R&D expenditures and capital investments in U.S. manufacturing. Section 3 sets up the model and defines the equilibrium concept. Section 4 discusses the assumptions required for the existence of such equilibrium and offers a constructive proof of its existence. Section 5 elaborates on the implications of the model towards the pattern of entry and exit, prices, sales and, in particular, the time-path of the investment and R&D series. A brief summary of the available empirical evidence on these aspects is presented. Conclusions close the paper.

2. The Industry Data

This section presents those features of the data that motivate the theoretical model. The data set extends over the period 1958-1983 and over 20 industry groups at the two and three digit level of the Standard Industrial Classification. The 20 industries and their SIC's are presented in Table 1 (see next page). The R&D data include privately and federally funded expenditures and comes from the Annual Surveys of Industrial R&D published by the National Science Foundation. The sales data are the value of shipments not adjusted for inventory changes while the investment data refers to new capital spending on structures and equipment. Both sales and investment come from the Annual Survey of Manufactures, and cover both R&D and non-R&D doers. All data was converted to 1982 dollars using the GNP deflator.[2]

Table 1 presents a succinct summary of the variables of interest: R&D and investment. The statistics are computed for the 26 years from 1958 till 1983. It reveals that the more R&D intensive industries, in terms of R&D expenditures per dollar of sales (R/S), are also the industries with the highest R&D-Investment (R/I) ratio. These industries are Aircraft and Missiles, Computers and Electrical Equipment, and Drugs. The correlation between R/S and R/I, across all industries, is 0.89. What may be surprising is that these industries spent more on R&D than on capital investment. In fact, seven industries spent more dollars on R&D projects than on plant and equipment over the sample period. Overall, the manufacturing sector spent 1.4 dollars in R&D per dollar of capital investment (0.93 dollars excluding the Aircraft and Missiles). For company financed R&D, this ratio is halved because of the important role played by

---

[2]We thank Sam Kortum for kindly providing the data set. A thorough explanation of the data is provided in Kortum (1990).

federally supported R&D in the Aircraft and Missiles industry. R/I declines to 0.69 (0.63 excluding Aircraft and Missiles). After 1974, there is a slight decrease in this ratio to 1.15, but it is still higher than unity (0.77 excluding Aircraft and Missiles).

Table 1: INDUSTRY AVERAGES, U.S. MANUFACTURING 1958-1983

INDUSTRY	SIC	R&D*	Invest*	R&D** Growth Rate	Invest** Growth Rate	R&D/*** Invest	R&D/*** Sales
1. Food and Kindred Products	20	532.47	5056.27	0.04	0.03	10.47	0.21
2. Textile & Apparel	22,23	124.28	2309.17	0.01	0.03	5.63	0.11
3. Lumber & Furniture	24,25	102.00	2113.06	0.07	0.02	4.51	0.15
4. Paper	26	385.34	3734.30	0.05	0.04	10.24	0.58
5. Industrial Chemicals	281-282,286	2480.73	5202.95	0.02	0.03	52.33	4.03
6. Drugs	283	1298.61	674.22	0.07	0.05	188.48	7.28
7. Other Chemicals	284-285,287-289	634.96	1657.56	0.03	0.04	42.83	1.34
8. Petroleum Refining & Extr.	29,13	1261.38	2957.50	0.04	0.05	47.80	1.48
9. Rubber & Plastics	30	638.55	1910.65	0.04	0.04	34.96	1.53
10.Stone, Clay & Glass	32	384.32	2369.17	0.03	0.01	16.66	0.86
11.Primary Metals	33	691.39	5836.23	0.03	0.00	12.40	0.55
12.Fabricated Metals	34	529.82	3132.16	0.01	0.03	18.18	0.50
13.Office Computing	357	3110.39	990.94	0.05	0.10	408.19	15.81
14.Other Nonelect.	351-356, 358-359	1574.19	4034.01	0.03	0.04	42.65	1.30
15.Communic.& Elect. Equip.	366-367	5794.73	2239.72	0.03	0.07	311.18	12.06
16.Other Elect.	361-365, 368-369	3775.62	1662.90	0.00	0.04	255.73	6.12
17.Trans. Equip.	371, 373-375, 379	4217.99	3876.22	0.03	0.05	127.78	2.92
18.Aircraft & Missiles	372, 376	12276.13	1391.46	0.01	0.03	1042.26	21.66
19.Instruments	38	1982.78	1179.42	0.06	0.06	165.83	5.53
20.Other Manufacturing	21,27,31,39	326.91	3169.22	0.01	0.03	10.61	0.28
ALL		2106.13	2774.86	0.03	0.04	140.44	4.22

Notes: * Millions of 1982 dollars.
** 3 year moving averages.
*** In percentage terms.

Over the sample period, the leading industry in terms of the level of R&D expenditures is Aircraft and Missiles. Its' mean R&D, 12 billion 1982 dollars per year, is about twice the mean R&D level of the second largest industry, the

Communication and Electrical Equipment industry. It is not, however, the largest growing industry. Among the industries increasing their annual expenditures on R&D by 5 percent or more we find some surprises such as Paper along with Drugs, Scientific Instruments, and Office Computing.[3] Computers is also the industry that is investing in plants and equipment at a much faster pace than the rest, 10 percent yearly, followed by the Communication and Electrical Equipment industry (7 percent yearly).

Over all industries and over the whole period 1958-1983, the levels of R&D and investment do not differ much. However, these are not evenly distributed across industries. R&D expenditures are as high (or even higher) than capital expenditures in most high-tech industries while considerable less so in traditional sectors. The inescapable conclusion is that, compared to capital investment, R&D is certainly not a negligible activity in the US manufacturing sector.

It is well known that investment at the aggregate level is very volatile relative to other aggregates such as GNP and consumption. The same is true within the U.S. manufacturing sector. Table 2 shows that the investment's volatility, as measured by the standard deviation of its 3 year moving average growth rate, is higher than R&D's in all but two industries. The same relationship holds in levels: the coefficient of variation of investment is higher than that of sales in 19 industrial groupings, and higher than that of R&D in 16.[4]

[3]In 1970 the Lumber and Furniture industry experienced an almost 200 percent increase in its R&D expenditures, from 18 to 52 million (nominal) dollars, which is responsible for the reported 7 percent annual increase in R&D.

[4]The almost 200 percent increase in R&D expenditures experienced by the Lumber and Furniture industry in 1970 is responsible for the large variability of R&D over time. The coefficient of variation up to 1969 is 17.6 (against 20.2

Table 2: VOLATILITY STATISTICS. U.S. MANUFACTURING 1958-1983

INDUSTRY	R&D* Growth Rate STD	Invest* Growth Rate STD	Sales** CV	Invest** CV	R&D** CV
1. Food and Kindred Products	0.03	0.03	0.14	0.22	0.27
2. Textile & Apparel	0.06	0.08	0.08	0.26	0.16
3. Lumber & Furniture	0.13	0.09	0.23	0.32	0.62
4. Paper	0.05	0.09	0.21	0.33	0.37
5. Industrial Chemicals	0.04	0.10	0.32	0.33	0.12
6. Drugs	0.02	0.08	0.30	0.42	0.50
7. Other Chemicals	0.07	0.10	0.24	0.39	0.22
8. Petroleum Refining & Extr	0.05	0.11	0.62	0.46	0.29
9. Rubber & Plastics	0.05	0.09	0.29	0.34	0.28
10.Stone, Clay & Glass	0.02	0.09	0.15	0.23	0.21
11.Primary Metals	0.04	0.11	0.19	0.24	0.22
12.Fabricated Metals	0.03	0.08	0.21	0.29	0.15
13.Office Computing	0.03	0.12	0.52	0.79	0.41
14.Other Nonelect.	0.04	0.11	0.27	0.40	0.28
15.Communic.& Elect. Equip.	0.06	0.10	0.32	0.59	0.21
16.Other Elect.	0.04	0.09	0.17	0.33	0.07
17.Trans. Equip.	0.04	0.14	0.21	0.53	0.23
18.Aircraft & Missiles	0.07	0.20	0.15	0.46	0.19
19.Instruments	0.06	0.08	0.33	0.42	0.50
20.Other Manufacturing	0.10	0.05	0.12	0.28	0.28

Notes: * Standard deviation of 3 year moving averages growth rates.
** Coefficient of Variation of the level of the variable (not in logs). CV = Standard Deviation/Mean.

Perhaps the easier way to form an impression of the data is to plot the investment and R&D series against time. Figure 1 presents such plots for all 20 industries. The variables being plotted are 3 year moving averages of the growth rates of investment and R&D to smooth out business cycle and size effects. Visual inspection indicates an investment series which is much more volatile than the R&D series in almost all industries.

---

for investment), and after 1971 it is 11.5 (against 19.3 for investment).

FIGURE 1: PLOTS OF R&D (□) AND INVESTMENT (–) GROWTH RATES, 1959-1983
3 - Year Moving Averages

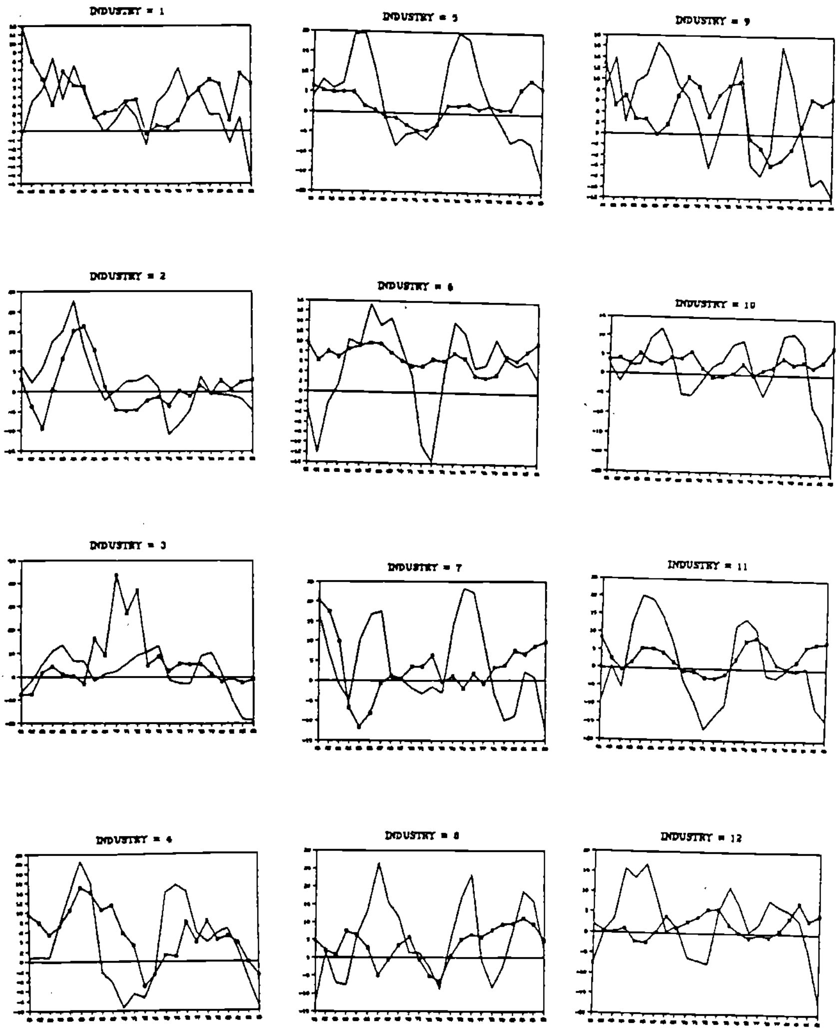
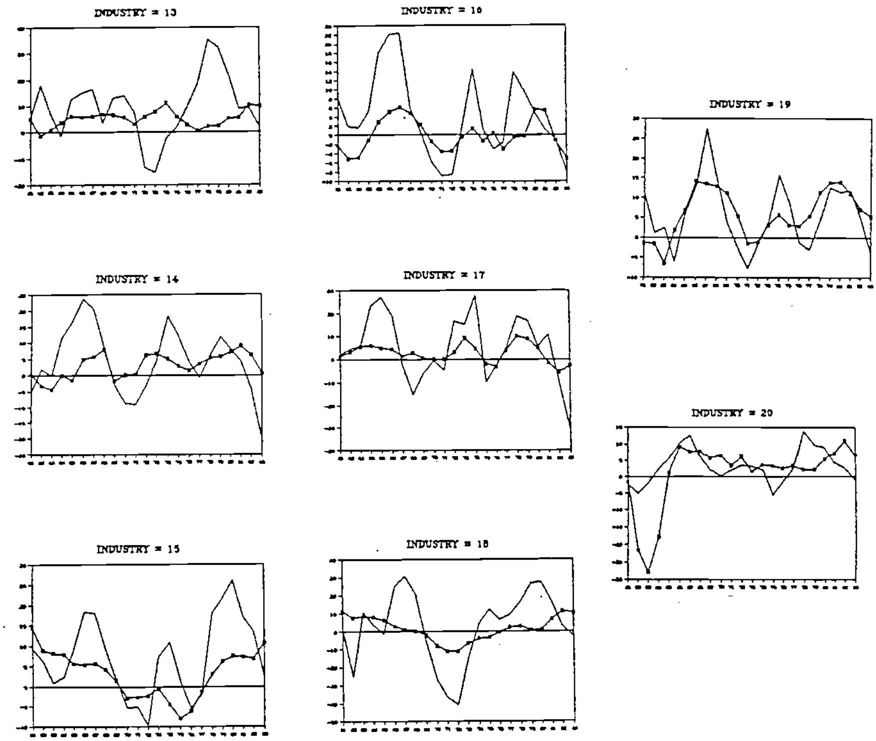


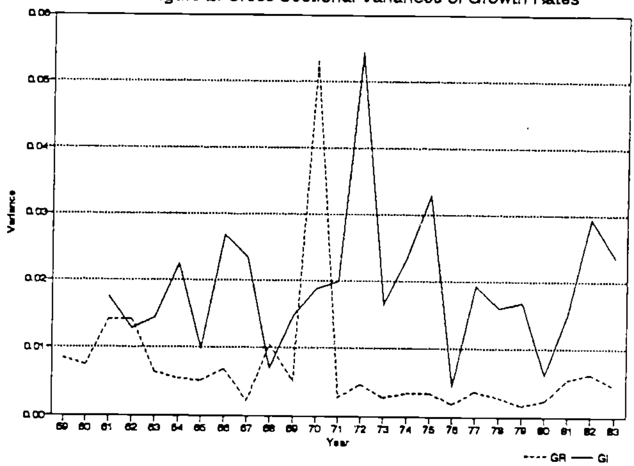
FIGURE 1: PLOTS OF R&D (□) AND INVESTMENT (–) GROWTH RATES, 1959-1983
3 - Year Moving Averages



These plots capture the variability of the series over time. Cross-sectional variances are also informative since they neutralize cyclical fluctuations and other macroeconomics events[5]. As can be seen in Figure 2, the cross-sectional variances of the R&D and investment growth rates obey the same relationship. Figure 2 also indicates that these variances remain, by and large, stationary over time.[6]

Figure 2

Figure 2: Cross-Sectional Variances of Growth Rates



The varinaces of GI in 1959 and 1960 are not shown. They equal 0.09 in both years.

---

[5]The total variances of R&D and capital investment growth rates are 0.0085 and 0.0350 respectively. The within-year (cross-sectional) variances account for 91 and 73 percent of them. In the other dimension, the within-industry variances account for 96 and 99 percent of the variances of R&D and investment growth rates.

[6]The outlier in the R&D variance is caused by the Lumber and Furniture industry. Deleting this industry decreases the 1970 variance form 0.053 to 0.0078. The first two outliers in the investment variance are due to the 90 percent increase in investment in 1959 and the subsequent 120 percent decrease in the Aircraft and Missiles industry. Deleting this industry decreases the 1959 and 1960 variances form 0.89 to 0.052 and to 0.028, respectively.

Finally, these features are not exclusive to industry-level data. Lach and Schankerman (1989) report a similar volatility pattern at the firm-level in science-based industries. In their sample of 191 firms, the sample variance of the logarithm of investment is 2.3 times higher than the sample variance of the logarithm of R&D expenditures, all variables measured as deviations from firms' means. The corresponding result for growth rates is 5.1 times. For only seven of the 191 firms is the ranking of these variances reversed. Similar results have been observed in other samples (Mairesse and Siu, (1984)).

In the next two sections we suggest a theoretical model which generates this excess volatility of investments. We then empirically verify the hypothesis underlying the model; namely, the causality relationship between R&D and investments.

3. The Model

3.1 An Informal Description of the Model

We analyze the dynamics of an industry in which a single, homogenous product is traded. Demand is constant[7] and the resulting dynamics are generated by an endogenous process of technological improvements being introduced into the industry by successive generations of firms. These improvements are such that the costs of production (both fixed and variable) are monotonically reduced over time.[8]

Technological improvements occur as a result of investments in R&D. The effect of these investments is to determine the distribution of the success date, i.e., the random date at which the R&D project is successfully completed and a

[7]For a formulation based on the uncertainty of demand see Rob (1991).

[8]For a related formulation see Aghion and Howitt (1990), who treat the successive monopoly case (as opposed to the oligopoly formulation here). Their model abstracts from capital-theoretic issues.

new innovation appears on the scene, ready to be implemented. The act of innovation is taken by an "R&D specialist firm", who then patents the new technology and sells it to a lone implementor in the production sector. The implementor acquires and installs machines, starting production and selling his output in a product-market where he competes with previous implementors (and future ones as they enter the market). New innovations are assumed to be always sold to new entrants, not to incumbent producers. This reflects some diseconomy in operating with multiple technologies. (The source of such diseconomies is not an integral part of the formulation).

To an outside observer the dynamics of the industry are as follows. At any given point in time the set of firms who potentially partake in the product market is historically given (those are the firms who had implemented the new technologies in the past). The product market is a Cournot oligopoly. The equilibrium outcome in it determines a market price, individual quantities, and a set of active producers, i.e., those with unit variable cost of production below the market price. As time progresses further innovations appear on the scene, and are implemented by newcomers. This reduces the market price and may trigger the exit of additional firms. This process of innovation, entry and induced exit takes place during the industry's growth phase. Eventually, given a fixed demand function, the technologies that are already in existence are so advanced, that the introduction of newer technologies is no longer profitable. New entry and R&D activity both cease at that point and the industry reaches a steady-state. The purpose of the model is to characterize the evolution of the endogenous data along such path.

3.2 Setting up the Model

A. The R&D Sector

Time is continuous, and is indexed by $t \in [0,\infty)$. The per-period interest rate is $r>0$. Demand is linear,

$$P = D(Q) = 1-Q, \qquad (1.1)$$

where Q is industry output and P is the price paid by consumers. Demand is assumed to be constant over time, and consumers are assumed to behave non-strategically.

At $t=0$, no firms are active in the market. Entry takes place at discrete points of time: at $t_1>0$ the first firm enters; at t_1+t_2 the second enters, and so on. Entry is always associated with the discovery of a new technology (an innovation or, more generally, a new idea). Hence, t_i represents the length of the i^{th} "discovery period". The time interval $\left[\sum_1^i t_j \, , \, \sum_1^{i+1} t_j\right]$ is referred to as "stage i".

When a firm enters, say firm i at $\sum_{j=1}^{i} t_j$ it implements a new technology. It does so by choosing a capacity level K_i and paying $m_i K_i$ dollars (up front). A unit of capacity is interpreted as the ability to produce one unit of output per-period, forever. We assume that capacity does not depreciate (alternatively, the cost of maintaining "machines" is included already in the variable cost) and that it can be increased later on (although this does not happen in equilibrium). Capacity cannot be liquidated, i.e., a firm cannot recover the historical cost, $m_i K_i$, by selling capacity. The extent to which capacity is utilized in subsequent periods is endogenous (see below).

The technology introduced at $\sum_{j=1}^{i} t_j$ is such that the per-unit capacity cost is $m_i \geq 0$, whereas the per-unit variable cost is $0 \leq c_i \leq 1$. Initially, we only assume that technological progress is characterized by:

$$m_1 > m_2 > \ldots$$
$$c_1 > c_2 > \ldots \qquad (1.2)$$

Further restriction on the sequences $\{c_i\}_{i=1}^{\infty}$, $\{m_i\}_{i=1}^{\infty}$ are imposed below so as to limit the range of equilibrium outcomes to those that are empirically relevant (see section 3.2). These sequences are exogenously fixed. Thus, the only uncertainty in the model stems from the fact that the sequence of discovery dates of new technologies, $\{t_i\}_{i=1}^{\infty}$, is not known in advance (see Aghion and Howitt (1990)).

The length of a typical discovery period, t_i, is stochastic and its distribution depends upon R&D expenditures during stage i-1. We denote these expenditures by $0 \leq x_{i-1} < \infty$, and consider a monopolistic (innovation) "producer" who chooses x_i to maximize her profit. The R&D sector is monopolized as a consequence, perhaps, of large fixed costs which turn this sector into a natural monopoly. The variable x_i is a flow variable that generates an instantaneous probability of discovery at instant t, given no discovery up to t, equal to:

$$h(x_i) \qquad (1.3)$$

Thus, the production technology in the R&D sector is memoryless in a broad sense: the probability of immediate success, given no success up to now, depends only on the current level of x and not on its past values nor on the type of discovery one is aiming for. This aspect of the R&D technology is standard in the patent-rate literature, and is represented by an exponential distribution function:

$$\Pr(\tilde{t} \leq t | x) = 1 - e^{h(x)t} \qquad (1.4)$$

where $\tilde{t}$ is a (random) discovery date and x is a flow of R&D expenditures. The following restrictions are placed on $h(\cdot)$:

$$h'(\cdot) > 0, \quad h''(\cdot) < 0, \quad h(0) = 0, \quad h'(0) < \infty \qquad (1.5)$$

Once a discovery comes about, it is sold to a firm which then enters into the product market, implementing the new technology. The price at which the discovery, or patent, is sold is denoted by V. V reflects the discounted value of future profits to the new entrant; this value is endogenous (see below). The net (of R&D expenditures) payoff to the monopolist who expends x dollars per unit time and expects to obtain V dollars upon discovery is:

$$R(x;V) = \int_0^\infty [Ve^{-rt} - x\int_0^t e^{-r\tau}d\tau]\, h(x)\, e^{-h(x)t}dt = \frac{h(x)V-x}{r+h(x)}. \qquad (1.6)$$

Maximizing the monopolist's payoff, the following first-order condition is derived:

$$[h'(x)V-1][r+h(x)] = h'(x)[h(x)V-x] \qquad (1.7)$$

Denote the (unique) x which solves this equation by $x^*(V)$. The following properties are straightforwardly derived:

$$x^*(V) > 0 \text{ if and only if } V > \underline{V} \equiv \frac{1}{h'(0)}, \qquad (1.8)$$

and

$$\frac{dx^*}{dV} = \frac{-h'(x)r}{h''(x)[x+Vr]} > 0 . \qquad (1.9)$$

Note that V can never exceed 1/4r, which is the <u>monopoly</u> profit under the demand curve (1.1), assuming zero cost of production (c=m=0). Thus, (1.9) implies the existence of maximal R&D expenditures, denoted by $\bar{x}$, and a maximal value for the hazard rate of discovery, denoted by $\bar{h}$. We summarize this in:

$$V \le \overline{V} \equiv \frac{1}{4r}, \quad x \le \overline{x}, \quad h \le \overline{h} \equiv h(\overline{x}) . \tag{1.10}$$

Finally, we assume that the industry is viable. That is, that profits to the first implementator under the best scenario possible (i.e., when she maintains a monopolistic position forever) are large enough to justify initial R&D expenditure. This is ensured by:

$$\frac{1}{r}\left[\frac{1+c_1+m_1 r}{2} - (c_1+m_1 r)\right]^2 > \underline{V} . \tag{1.11}$$

B. The Product Market

The product market is modelled as a Cournot oligopoly, quantities being the strategic variables. Consider a point in time, $\sum_{j=1}^{i} t_j$, at which firm i is to make a capacity decision (K_i), implementing the i^{th} technology, (c_i,m_i). At that point an asymmetry arises: firm i must pay capacity costs ($m_i K_i$), whereas incumbent firms need not. This asymmetry is based on the presumption that firm i's predecessors, firms i-1,...,1, are already in the market, and are planning to produce from this point onwards below their pre-installed capacity. In section 3.2 we impose sufficient conditions on the fundamentals of the model to ensure that this presumption is indeed verified.

Denote the quantities produced by different vintage firms during stage i by $q_{i,j}<K_j$, $1 \le j<i$, and $q_{i,i}=K_i$. Hence, one can dispense with the variable K_i, identifying it with $q_{i,i}$.

R&D expenditures during stage i, x_i, are determined by the monopolistic R&D

sector and are independent of what occurs in the industry during that stage.[9] Hence, each firm takes x_i, and its implications on the length of stage i, as exogenously fixed. Firm i's objective then is to maximize:

$$\int_0^\infty\int_0^t q_{i,i}\left[1-\sum_{k=1}^{i} q_{i,k}-c_i\right]e^{-r\tau}d\tau\; h(x_i)\,e^{-h(x_i)t}dt - m_i q_{i,i} = \tag{2.1}$$

$$= \frac{q_{i,i}\left[1-\sum_{k=1}^{i} q_{i,k}-c_i\right]}{r+h(x_i)} - m_i q_{i,i}$$

whereas firm j's objective is to maximize:

$$\frac{q_{i,j}\left[1-\sum_{k=1}^{i} q_{i,k}-c_j\right]}{r+h(x_i)}\ , \quad i\le j<i\ . \tag{2.2}$$

Expressions (2.1) and (2.2) describe the expected discounted profits during stage i. Discounting is taken with respect to the factor e^{-r} whereas expectations are taken with respect to the length of stage i which, as indicated above, is exponentially distributed with parameter $h(x_i)$. To abbreviate, we shall use the notation $h_i \equiv h(x_i)$.

The Cournot equilibrium of the game defined by the above objective functions is as follows. There exists an integer L_i, $0\le L_i<i$, denoting the "last firm to exit", such that:

---

[9] The monopolist only cares about the price, V_{i+1}, at which she sells the forthcoming patent. This price depends only on the industry's evolution from $\sum_1^i t_j$ onwards, and not on the industry's evolution from outcomes of prior stages. These assertions are substantiated in the proposition below.

$$P_i = \frac{1+c_i+m_i(r+h_i) + \sum_{k=L_i+1}^{i-1} c_k}{i-L_i+1}\,, \tag{2.3}$$

where P_i denotes the equilibrium product price during stage i. And,

$$c_i + m_i(r+h_i)\,,\ c_{L_i+1} < P_i < c_{L_i}. \tag{2.4}$$

These inequalities reflect the idea that firm L_i and all its predecessors have exited, their variable costs being above the product price, while firm L_i+1 and all its successors are active producers. Hence, condition (2.4) describes the pattern of exit from the industry. Next,

$$q_{i,j} = \begin{cases} 0 & j=1,2,\ldots L_i \\ P_i-c_j & L_i+1 \le j < i \\ P_i-[c_i+m_i(r+h_i)] & j=i \\ 0 & j>i \end{cases} \tag{2.5}$$

describes the (Cournot-equilibrium) quantities produced by the various firms at the various stages. And,

$$\pi_{i,j} = q_{i,j}^2\,, \tag{2.6}$$

describes the flow profits to each firm at each stage.

C. The Equilibrium Concept

Using the above notation we may write firm i's discounted profit as follows:

$$V_i = \sum_{k=0}^{\infty} \frac{\pi_{i+k,i}}{r+h_{i+k}} \prod_{\tau=0}^{k-1} \frac{h_{i+\tau}}{r+h_{i+\tau}} = \sum_{k=i}^{\infty} \frac{\pi_{k,i}}{r+h_k} \prod_{\tau=i}^{k-1} \frac{h_\tau}{r+h_\tau} \tag{2.7}$$

The terms in the above expression(s) represent firm i's profit over the different stages of its life-cycle — discounted to its point of inception i.e., to $\sum_{j=i}^{i} t_j V_i$ represents the (stock market) value of firm i once its optimal capacity, $q_{i,i}$, is installed and paid for. Assuming perfect competition among potential buyers of patent i, V_i is the price at which patent i is sold to a new entrant.

We can now introduce an equilibrium concept, reflecting the interdependencies among the three economic activities in the model; namely, R&D, physical investments and per-period production. Therefore, the equilibrium is such that the levels of these activities are simultaneously determined.

Definition: An equilibrium consists of sequences

$$\{(V_i),\ (x_i, h_i),\ (P_i),\ (L_i),\ (q_{i,j}),\ (\pi_{i,j})\}_{i,j=1}^{\infty}\ ,$$

so that the following requirements hold:

$h_i = h(x_i)$

$x_i = x^*(V_{i+1})$

P_i satisfies (2.3)

L_i satisfies (2.4)

$q_{i,j}$ satisfies (2.5)

$\pi_{i,j}$ satisfies (2.6)

and V_i satisfies (2.7), $i,j = 1,2,\ldots.$

The first two requirements in the above definition reflect maximization in the R&D sector, the next four reflect equilibrium in physical investments and in the product market, whereas the last requirement links together the various variables in the model.

4. Analysis

In this section we construct an equilibrium, proving thereby its existence. The construction hinges on several assumptions, the effect of which is to restrict the values of the model's parameters to lie within certain regions. The following remarks briefly explain the purpose of these assumptions. They also suggest the type of industries where these assumptions are likely to be satisfied.

The reason for introducing additional restrictions lies in the existence of externalities (or "spillovers") in the R&D sector. These externalities are due to the fact that in doing R&D one is building on the base of knowledge created by one's predecessors, but one is not remunerating those predecessors. As a consequence, the incentives to engage in R&D are blunted, which may result in a degenerate equilibrium , i.e., in a situation in which the industry never opens. This possibility is illustrated by a simple example contained in the appendix. Thus, without any parametric restrictions, the theoretical possibility arises where an evolutionary growth-path is either "abrupt" or simply non-existent.

On the other hand, common sense and the data we report below suggests the existence of "diffused" growth paths; in fact, the essence of our empirical findings is to verify the behavior of certain variables along such time-paths. Accordingly, the restrictions stated in the proposition below limit the theoretical scope of equilibrium configurations. They do this by placing a limit on the effect of R&D spillovers, which guarantees that the private return to R&D is sufficiently high. Thereby, activity in the R&D sector is not stifled, and the industry's growth path is diffused, not abrupt.

A special case of production processes which satisfies the set of assumptions stated below is where (a) technical progress affects primarily the capital component of production, i.e., where m_i is reduced much faster than c_i; and where (b) the capital component accounts for the bulk of production expenses, i.e., where $m_i r >> c_i$.[10] (Clearly, these two conditions are not independent: R&D efforts are likely to be concentrated on the costlier component of production). Case (a) is specially appealing given the extraordinarily rapid pace of advance in electronics process technology. As Rosenberg and Steinmuller (1982, page 181) put it: "New IC (integrated -circuit) products tend initially to be both more expensive and of higher performance. But, subsequently, the price of these higher-performance devices falls so that the all important cost per unit of performance dramatically declines as each new product reaches the mature stage of the product cycle".[11] In particular, the costs associated with computerization of production and administrative systems has been greatly reduced.[12] Case (b) is characteristic of industries such as drugs, chemicals, transportation, plastics, and machinery where the process of production requires investments in heavy equipment.

Proposition:

Assume that for all $i \geq 2$:

(i) $c_i + m_i r > c_{i-1}$

---

[10] The most extreme case of this is where $c_i = c$, and $m_i r > c$, for all i.

[11] In the same article the authors quote from A. Osborne's book (Running Wild - The Next Industrial Revolution): "...a ticket for a Concorde flight would have to cost less than a penny if it were to compare with the rate at which microelectronics has gotten cheaper".

[12] For example, the decline in the real price of personal computers, adjusted for quality improvements, is estimated to be around 25 percent annually during the 1980's (see Berndt and Griliches (1990)).

(ii) $c_{i-1} - c_i < \varepsilon$, where ε is defined below

(iii) $c_i < \frac{c_{i-1}+m_{i-1}r}{2}$

(iv) $m_i(r+\bar{h}) \leq \frac{1}{(i+1)(i+2)}$

(v) For every i, there exists a j_i, $0 \leq j_i \leq i-2$, so that:

$$c_{j_i} > \frac{1+m_i r+c_i+\ldots+c_{j_i+1}}{i-j_i+1} > c_{j_i+1},\ c_i+m_i r$$

Then there exists an equilibrium where each firm produces below its preinstalled capacity in all stages except for the stage at which it first enters into the market.

Proof: The proof is simple, but long. We relegate it to appendix A.

Remarks: 1. The special case alluded to above is where $c_i = c$, where $m_i r > c$ and where only condition (iv) is required to hold. As one can verify, all other restrictions are automatically satisfied in this case.

2. Condition (iv) states that innovations are sufficiently drastic (see Arrow (1962) for the origins of this condition), i.e., that m_i diminishes rapidly enough. This condition ensures the profitability of new entrants and, at the same time, it ensures that incumbents are producing below their preinstalled capacity.

5. Implications and Empirical Results

5.1 Industry Dynamics

The construction and analysis of the model point towards an industry that, along its equilibrium path, sees its number of firms increase at first, followed by simultaneous entry and exit. Total industry output increases and prices decline. Nevertheless, each firm, leaders and incumbents alike, produces less output and has lower current profit as time goes on. Furthermore, in the special case (see remark 1 above) where $c_i = c$ additional monotonicity properties hold: prices decline and total output increases at a <u>declining</u> rate; the stock market values of entering firms, V_i, and the rate of technological improvements, h_i, both decrease over time.

Broadly speaking, these implications seem to be confirmed by the data presented in such studies as Gort and Klepper (1982) and Klepper and Graddy (1990). They analyze the evolution of 46 new products, which include consumer and producer goods, from the date of initial commercial introduction of the product through 1972. Among their many interesting observations, the ones relevant to the present paper are the following. Initially, the number of firms in the industry grows but, later on, industries experience a period in which there is a decline in the number of producers. This shakeout stage is followed by a stabilization in the number of firms. All along, there is a continuous process of gross entry into and gross exit from industry. During both the growth and the shakeout stages total industry output grows and price declines at a decreasing percentage rate. During the stage in which the number of firms stabilizes, the output grows and price falls at a constant percentage rate. Furthermore, data on counts of innovations suggest the "major innovations" occurring during the early stages of development of new product industries are of greater importance than those occurring later. This data, however, do not indicate a decline in the number of innovations over time; it seems to remain constant.[13] Finally, and most relevant to the model in this paper, there appears to be an association between rises and declines in the rate of innovation and the rate of entry into new markets. Gort and Klepper interpret the causal relation as being positive, and flowing primarily from innovations to entry rates during the period of positive net entry.

5.2 Investment and R&D

A central point of this paper is that each time a new discovery is made it generates a new burst of economic activity: entry and exit, additional investment and output. But, discoveries or innovations are not exogenous to the economy. Rather, they represent the uncertain outcomes of R&D activity. These outcomes, in turn, partly depend on the investments in R&D done by the firms. This reasoning establishes a causal nexus between R&D expenditures on the one hand and measures of economic activity on the other.

Lach and Schankerman (1989) present empirical evidence on this issue in their analysis of a panel of large firms in the high-tech sector of U.S. manufacturing during the 1970's. They find that R&D expenditures Granger cause investment in physical capital and not the other way around. Here we extend these findings to the industry-level case, using the same data set described in section 2.

A word of caution is in order regarding the empirical results. To properly test the implications of the theoretical model we would ideally like to know the

[13] Their patent data indicates an increase over time in the rate of patenting but it is well known that the patent series have a lot of noise and, as the authors conclude, are not a good measure of the rate of technological change. On the other hand, Kortum (1990) finds a large decline in patents per R&D during the period 1958-1983 in all of his 20 industrial groups (2-3 digit level). This decline is caused by reduction in the absolute number of patents granted to the industries.

timing of the innovations. This knowledge would enable us to compute the "correct" changes in investment and R&D as well as guide us in the analysis of the causality pattern. Needless to say, the industry level data do not provide this kind of information.[14] In addition, the observed data reflects the effects of many factors which were ignored in the theoretical model such as the (artificial) distinction between R&D doers and producers, business cycles effects and the impact of an evolving market size. Some of these issues have been taken care of in the empirical work, albeit in an ad-hoc fashion. The results are, therefore, meant to be suggestive and conducive to future research. The paper by Lach (1992) deals more systematically with these issues.

Table 3 presents the regression coefficients and, in the lower panel, the results of the tests for the direction of Granger causality. The first two columns present the results of OLS regressions of investment and R&D on 3 lags of both variables. The joint test for the absence of all three or the first two lags of R&D in the investment equation is strongly rejected. Conversely, the point estimates of the parameters of lagged investment in the R&D equation are very close to zero, and the hypothesis of no effect of past investment on R&D cannot be rejected. Columns 3 and 4 report similar regressions when, in addition to lagged R&D and investment, lagged sales are added to each equation. The estimates do not appear to change much because of these additions. Finally, the last two columns perform the causality tests on the growth rates of R&D and investment yielding similar results.

---

[14]This ideal situation could probably be well approximated by "case-study" type of data sets and, in particular, case-studies of those firms which are involved in big R&D projects. An example may be the data set of the Bell System analyzed in Gordon, Schankerman, and Spady (1986).

Table 3: CAUSALITY TESTS. U.S. MANUFACTURING, 1958–1983[a]

	I (1)	R (2)	I (3)	R (4)	g^{I} [b] (5)	g^{R} [b] (6)
I_{-1}	0.92	0.01	0.78	−0.03	0.06	0.02
	0.06	0.02	0.07	0.03	0.06	0.02
I_{-2}	−0.15	0.00	−0.12	0.01	−0.17	−0.00
	0.08	0.03	0.08	0.03	0.04	0.02
I_{-3}	0.02	−0.00	0.04	0.00	−0.05	−0.00
	0.05	0.03	0.05	0.03	0.04	0.02
R_{-1}	0.28	0.96	0.21	0.94	0.24	0.05
	0.07	0.08	0.06	0.09	0.08	0.08
R_{-2}	−.021	0.11	−0.14	0.12	0.02	0.13
	0.09	0.15	0.08	0.15	0.08	0.11
R_{-3}	−0.03	−0.21	−0.03	−0.20	−0.06	−0.11
	0.07	0.11	0.07	0.11	0.07	0.05
S_{-1}		0.64	0.16			
		0.18	0.06			
S_{-2}		−0.37	−0.07			
		0.20	0.06			
S_{-3}		−0.13	−0.04			
		0.13	0.05			
N	460	460	460	460	440	440
Adj. R^2	0.96	0.99	0.96	0.99	0.34	0.06
Serial Correlation[c] (p–value)	0.01	0.10	0.01	0.11	0.12	0.05

PROBABILITY VALUES OF CAUSALITY TESTS[d]

	I (1)	R (2)	I (3)	R (4)	g^{I} (5)	g^{R} (6)
H_0						
$R_{-1}=R_{-2}=R_{-3}=0$	0.001		0.013		0.029	
$R_{-1}=R_{-2}=0$	0.000		0.005		0.012	
$I_{-1}=I_{-2}=I_{-3}=0$		0.826		0.755		0.938
$I_{-1}=I_{-2}=0$		0.725		0.566		0.852

[a] Industry and year dummies are included in all regressions. Heteroskedastic consistent standard errors appear at the bottom line of each entry.

[b] Growth rates are regressed on lagged growth rates, not on lagged levels.

[c] Probability values of LM test for first order serial correlation.

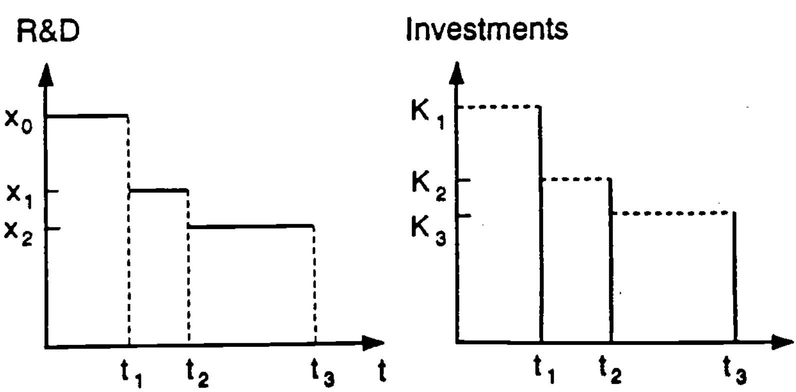
[d] H_0 is rejected at the α percent significance level whenever the p–value is less than $\alpha/100$.

The picture that emerges is that past R&D matters in the investment equation but past investment does not affect current R&D. Or, R&D Granger-causes investment but investment does not (Granger) cause R&D. These results support the assumption that lies behind the theoretical model and is consistent with similar results found at firm-level data.

Another manifestation of the causal relationship between R&D and investment in physical capital is that it provides an alternative explanation for the higher volatility of the investment time series relative to that of the R&D time series, as per the evidence presented in section 2. This empirical observation is also implied by our model of industry dynamics with respect to investment and R&D expenditures at the industry level.

The equilibrium of our model is characterized by both the R&D and physical capital series declining over time. However, the former is a smooth series with downward steps at discrete points in time, while the latter consists of spikes at the same points in time, also declining in magnitude over time — see Figure 3.

Figure 3



This pattern is a result of the distinct nature of investment in knowledge (R&D) and investment in physical capital. The latter responds to the realized values of the R&D process, i.e. to ideas for new products or production processes, and comes about in the form of additional purchases of new equipment and structures necessary for their implementation, i.e., additional capital investment. On the other hand, the response of R&D resources to R&D outcomes is of a different nature. In cases where the distribution of R&D outcomes or completion times are known with certainty, as in our model, observed values of the R&D process do not provide additional information to the firm and, therefore, do not affect subsequent R&D investments. Hence, there is no variability of the R&D series on this account.[15] In other cases, the firm will update its prior beliefs on the distribution of R&D outcomes or completion times using the realized values of the R&D process. In general, though, it seems reasonable to expect R&D's response to R&D's observed outcomes to be much smoother than the response of physical investment; the channels by which these responses are elicited being of a different nature for each type of investment.[16]

In sum, the theoretical model explicitly posits a causality pattern between the two investment activities, from R&D to physical capital investment. Such pattern, coupled with the uncertainty in the timing of the innovations, implies a more volatile capital expenditures series at the industry level. The model's implications on causality and volatility are consistent with the data.

---

[15] Of course, both types of investment respond to market conditions; probably in different ways.

[16] For an alternative explanation of the excess variability of investment relative to R&D see Hall and Hayashi (1989).

6. Conclusions

It is widely accepted that the arrival of innovations is one of the key determinants of industrial growth (see Schumpeter (1939) or Nelson and Winter (1982)). Innovations act like a trigger that sets in motion a set of activities such as entry into and exit from the industry, new investments, and changes in the levels of production. It is also accepted that these innovations are not exogenous to the economy. Rather, they are the fruit of a continuous process of investment in research and development performed by profit maximizing firms.

This paper presents a simple model of industry growth that captures this scenario. It emphasizes the causal relationship between expenditures in R&D, its output in the form of inventions, and subsequent expenditures necessary for the implementation of such inventions, such as the acquisition of new machinery and equipment, i.e., capital investment. The implications of the model regarding the evolution of the endogenous variables (output and prices, entry and exit rates) are consistent with the empirical observations of Gort and Klepper (1982) and Klepper and Graddy (1990). The model also predicts that, along the equilibrium path, capital investment exhibits greater variability than R&D expenditures. Data on twenty U.S. manufacturing industries for the period 1958–1983 confirm this prediction. Regression results indicate that, as assumed in the model, R&D (Granger) causes capital investment and not the other way around.

References

Aghion, P. and Howitt, P. "A Model of Growth through Creative Destruction" Working Paper, University of Western Ontario, April 1990.

Arrow, K., "Economic Welfare and the Allocation of Resources for Invention", in R. R. Nelson (ed.), The Rate and Direction of Inventive Activity: Economic and Social Factors. NBER Conference Series No 13, Princeton: Princeton University Press, 1962.

Berndt E. and Z. Griliches, "Price Indexes for Microcomputers: An Exploratory Study", NBER Working Paper No. 3378, June 1990.

Dixit, A., "The Role of Investment in Entry Deterrence", Economic Journal Vol 90, 1980, 95-106.

Ericson, R. and A. Pakes, "An Alternative Theory of Firm and Industry Dynamics", mimeo, Columbia University, 1990.

Flaherty, M.T. "Industry Structure and Cost reducing Investment", Econometrica 48 (1980), 1187-1210.

Fudenberg, D. and J. Tirole, "Dynamic Models of Oligopoly", Vol. 3 of Fundamentals of Pure and Applied Economics eds., J. Lesourne and H. Sonnenschein, London: Harwood, 1986.

Gilbert, R.J. and Newberg, D.M.G., "Preemptive Patenting and the Persistence of Monopoly", American Economic Review 72(1982), 514-526.

Gordon, R., Schankerman, M. and R.Spady, "Estimating the Effects of R&D on Bell System Productivity: A Model of Embodied Technological Change," in M. Peston and R. Quandt, eds., Prices, Competition and Equilibrium (Philip Allan Publ: Oxford, 1986), 164-188.

Gort, M. and S. Klepper, "Time Paths in the Diffusion of Product Innovations", Economic Journal, September 1982, 630-653.

Griliches, Z., ed., R&D, Patents and Productivity, Chicago: University of Chicago Press, 1984.

Hall, B. and F. Hayashi, "Research and Development as an Investment", NBER Working Paper No. 2973, 1989.

Harris, C. and Vickers, J. "Perfect Equilibrium in a Model of Race" Review of Economic Studies 52(1985), 193-209.

Jovanovic, B. and MacDonald G. "The Life-Cycle of a Competitive Industry: Theory and Measurement," Working paper No.237 Rochester University.

Kamien, M.I and Schwartz N.L., Market Structure and Innovation, Cambridge: Cambridge University Press (1980).

Klepper, S. and E. Graddy, "The Evolution of New Industries and the Determinants of Market Structure", The Rand Journal of Economics 21, 1990, 27-44.

Kortum, S., "A Structural Model of Inventions, R&D and Industry Growth", Yale University, 1990.

Loury, G. "Market Structure and Innovation," Quarterly Journal of Economics 93 (1979), 395-410.

Lach, S. "The Relationship between Research and Capital Investment", in Three Essays on the Economics of Research and Development and its Interaction with Capital Investment, Ph.D. Dissertation, Columbia University, 1988.

Lach, S. and M. Schankerman, "Dynamics of R&D and Investment in the Scientific Sector," Journal of Political Economy 97 (August 1989), 880-904.

Lach, S. "Investment in Knowledge and in Capital", Mimeo, Hebrew University Jerusalem.

Mairesse, J. and A.K. Siu, "An Extended Accelerator Model of R&D and Physical Investment," in Z. Griliches, ed., R&D, Patents and Productivity, Chicago: University of Chicago Press, 1984, 271-297.

Mueller, D.C., Profits in the Long-run, Cambridge: Cambridge University Press, 1986.

Nelson, R. and Winter, S. "An Evolutionary Theory of Economic Change", Harvard University Press: Cambridge, MA 1982.

Reinganum J. "Uncertain Innovation and the Persistence of Monopoly", American Economic Review 73(1983), 741-748.

Reinganum, J., "Practical Implications of Game Theoretic Models of R&D", American Economic Review, Papers and Proceedings, 74, 1984, 61-66.

Reinganum, J.,"Innovation and Industry Evolution", Quarterly Journal of Economics, 100, 1985, 81-100.

Rob, R., "Learning and Capacity Expansion in a New Market under Uncertainty", Review of Economic Studies 58 (1991), 655-675.

Rosenberg, N. and W. Steinmuller, "The Economic Implications of the VLSI Revolution" in Inside the Black Box, Technology and Economics, by Nathan Rosenberg, Cambridge University Press, 1982.

Schumpeter, J. The Theory of Economic Development. Cambridge, MA: Harvard University Press (1934).

Appendix

A. Proof of Proposition: Denote the middle term in condition (v) by z_i, and note that

$$c_\infty \le z_i \le \frac{1+m_i r+c_i+\ldots+c_1}{i+1}, \tag{3.1}$$

where $c_\infty = \lim_{i\to\infty} c_i$. The left hand inequality follows from the fact that z_i is an arithmetic average of terms that are greater than c_∞, while the right hand inequality follows from the fact that

$$\frac{1+m_i r+c_i+\ldots c_1}{i+1} = \frac{1+m_i r+c_i+\ldots c_{j_i+1}}{i-j_i+1}\,\frac{i-j_i+1}{i+1} + \frac{c_{j_i}+\ldots c_1}{j_i}\,\frac{j_i}{i+1}, \tag{3.2}$$

and from the definition of j_i which implies

$$\frac{c_{j_i}+\ldots+c_1}{j_i} > \frac{1+m_i r+c_i+\ldots+c_{j_i+1}}{i-j_i+1}. \tag{3.3}$$

Now, the right hand side of (3.1) converges to c_∞ (as a Cesaro average of a sequence that converges to c_∞). Thus, $\lim_{i\to\infty} \{z_i-[c_i+m_i r]\} = 0$. Hence, there exists an N such that:

$$V_N \equiv \frac{[z_N-(c_N+m_N r)]^2}{r} > \underline{V}, \tag{3.4}$$

whereas for all n>N:

$$\frac{[z_n-(c_n+m_n r)]^2}{r} \le \underline{V}\ . \tag{3.5}$$

From (1.11) it follows that $N\ge 2$.

We shall now use "backwards construction" to exhibit an equilibrium where firm N is the last to enter. Thus, stage N corresponds to the industry's "steady-state", its length being infinite (i.e., $x_N=h_N=0$). The product price during that stage is $P_N=z_N>c_N+rm_n$. Thus, firm N is profitable. Furthermore, firm N's value, V_N is sufficiently large (i.e. $V_N\ge\underline{V}$) to induce R&D effort leading to the N^{th} discovery. Note also, from condition (i), that firm N-1 is profitable during stage N, i.e., $c_{N-1}<c_N+m_Nr<P_N$ and $q_{N,N-1}>q_{N,N}$ (this presumes, though, that firm N-1 has preinstalled capacity, a presumption which we now confirm).

Define x_{N-1} as the maximizer of $R(x,V_N)$, and let $h_{N-1}=h(x_{N-1})$. Consider the (N-1) stage oligopoly, the set of potential producers in it being (1,2,...N-1). Let the set of actual producers be (m,...,N-1), with cardinality k(=N-m). Then

$$P_{N-1} = \frac{1+c_m+\ldots+c_{N-1}+m_{N-1}(r+h_{N-1})}{k+1}, \tag{3.6}$$

whereas,

$$P_N \le \frac{1+c_m+\ldots+c_{N-1}+c_N+m_{N-1}r}{k+2}. \tag{3.7}$$

The term on the right hand side of (3.7) presumes that no firm drops out between stages N-1 and N; since dropouts are always by inefficient firms, their effect is to <u>lower</u> prices. Hence, the right hand side is indeed an upper bound on stage N price. Combining (3.6) and (3.7) we have:

$$\begin{aligned} P_{N-1}-P_N &\ge \frac{1+c_m+\ldots+c_{N-1}+m_{N-1}(r+h_{N-1})}{k+1} - \frac{1+c_m+\ldots+c_{N-1}+c_N+m_Nr}{k+2} \\ &\ge \frac{1+c_m+\ldots+c_{N-1}-(k+1)\,c_N+m_{N-1}(r+h_{N-1})}{(k+1)\,(k+2)} \ge m_{N-1}(r+h_{N-1}), \end{aligned}$$

where the last inequality follows from the fact that c_m, $c_{m+1},\ldots,c_{N-2}\ge c_N$, from condition (iii) which implies $c_{N-1}+m_{N-1}(r+h_{N-1}) \ge 2c_N$, and from condition (iv).

Hence firm N-1 is profitable during stage N-1, i.e., $P_{N-1} > c_{N-1}+m_{N-1}(r+h_{N-1})$ (as remarked above, it is profitable during stage N; and stage N-1 price is higher than stage N price by more than its capacity installment cost, $m_{N-1}(r+h_{N-1})$.) Furthermore, $q_{N,N} < q_{N,N-1} < q_{N-1,N-1}$.

Let us now verify that firm N-1's value, V_{N-1}, is large enough to induce R&D expenditures aimed at attaining the $(N-1)^{th}$ innovation, i.e., let us verify that $V_{N-1} \ge \underline{V}$. By (2.7) above:

$$V_{N-1} = \frac{q^2_{N-1,N-1}}{r+h_{N-1}} + \frac{q^2_{N,N-1}}{r}\,\frac{h_{N-1}}{r+h_{N-1}} = \frac{r}{r+h_{N-1}}\,\frac{q^2_{N-1,N-1}}{r} + \frac{h_{N-1}}{r+h_{N-1}}\,\frac{q^2_{N,N-1}}{r} > \frac{q^2_{N,N}}{r} = V_N \geq \underline{V}\,,$$

since, as we have already remarked, $q_{N-1,N-1}$, $q_{N,N-1} > q_{N,N}$. Given V_{N-1} we define $x_{N-2}=x^*(V_{N-1})$ and $h_{N-2}=h(x_{N-2})$.

We now turn to the stage N–2 oligopoly where we must again verify: (a) that firm N–2 is profitable, $c_{N-2}+m_{N-2}(r+h_{N-2}) < P_{N-2}$ and (b) that firm N–2's value justifies its R&D expenditures, $V_{N-2} \geq \underline{V}$.

The proof of the first claim is completely analogous to firm (N–1)'s case, so it will not be repeated. As to the second claim we have:

$$V_{N-2} = \frac{r}{r+h_{N-2}}\,\frac{q^2_{N-2,N-2}}{r} + \frac{rh_{N-2}}{(r+h_{N-2})\,(r+h_{N-1})}\,\frac{q^2_{N-1,N-2}}{r} + $$

$$+ \frac{h_{N-2}h_{N-1}}{(r+h_{N-2})\,(r+h_{N-1})}\,\frac{q^2_{N,N-1}}{r}\,.$$

This convex combination is such that $q_{N-2,N-2} > q_{N-1,N-2} > q_{N-1,N-1}$, whereas for sufficiently small $\epsilon>0$ (Note that:

$$|\, q_{N,N-2} - q_{N,N-1} \,| = |\, P_N - c_{N-2} - (P_N - c_{N-1}) \,| = |\, c_{N-1} - c_{N-2} \,|)\,,$$

$$|\, c_{N-1} - c_{N-2} \,| < \epsilon \rightarrow \left| \frac{q_{N,N-2}}{r} - \frac{q_{N,N-1}}{r} \right| < \delta \rightarrow V_{N-2} \geq \underline{V} \text{ (or even } V_{N-2} \geq V_{N-1} \text{) }.$$

This process can be continued all the way to firm 1, yielding the $\epsilon>0$ stated in the proposition. The proof is complete.

B. An example where R&D spillovers are sufficiently strong to cause a degenerate equilibrium.

We let $m_1 = m_2 = \ldots = 0$, $c_1 = 0.6$, $c_2 = c_3 = \ldots = 0$, $r = 0.1$, and $h'(0)=1$ (eg, $h(x) = \frac{x}{x+1}$).

Then $c_1 > c_2 + m_2 r$. Thus, the second stage entrant monopolizes the industry, her discounted profit being $V_2 = \frac{1}{4r} = \frac{1}{0.4} = 2.5$. This satisfies $V_2 > \frac{1}{h'(0)} = 1$, as per (1.8) above. Hence, there is an incentive to create the second innovation. The intensity of R&D leading to that innovation, ie, x_2, is determined from (1.7):

$$(\frac{1}{4} + x)h'(x) = r + h(x) \qquad (1.7')$$

Denote the solution to this by x_2, and let $h_2 = h(x_2) > 0$. The prize facing the first innovator is then

$$V_1 = (\frac{1-c_1}{2})^2 \frac{1}{r+h_2} \leq 0.4,$$

substituting for c_1, and noting that V_1 is diminishing in h_2. Therefore, according to (1.8), $x_1^* = 0$ and the industry never opens. The example is somewhat extreme in that it assumes $c_2 = c_3 = \ldots = 0$. On the other hand, it is robust in that small perturbation of the data will leave the degeneracy of equilibrium intact.