

**NBER WORKING PAPER SERIES**

**RESEARCH AND PRODUCTIVITY**

**Boyan Jovanovic  
Yaw Nyarko**

**Working Paper 5321**

**NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 1995**

**We thank the C.V. Starr Center for Applied Economics for technical and financial help, Zvi Griliches and Bronwyn Hall for comments, and Chung Tse for helping with the research. 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.**

**© 1995 by Boyan Jovanovic and Yaw Nyarko. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.**

RESEARCH AND PRODUCTIVITY

**ABSTRACT**

We model research as a signal on an unknown parameter of a technology. We distinguish applied from basic research and show that firms in the same industry can optimally choose different research portfolios, and that basic research can seem to have a higher rate of return than applied research, even though it really doesn't -- essentially, firms on a "fast track" upgrading policy opt for basic research but fast and slow-track upgrading policies can coexist in a long-run equilibrium.

We also derive the lag structure for how R&D affects the firm's stock of knowledge. To a first approximation, the lags decay geometrically (as is typically assumed in practice) but the rate of decay is endogenous, and depends on how fast the firm is upgrading its technology.

Boyan Jovanovic  
Department of Economics  
University of Pennsylvania  
Philadelphia, PA 19104  
and NBER

Yaw Nyarko  
Department of Economics  
New York University  
New York, NY 10003

## 1. Introduction

We propose an informational theory of how research raises productivity. We model research as a signal on an unknown parameter of a technology. Such a signal teaches the firm how to operate the technology it uses now, or one that it intends to use in the future. The paper deals with two topics. The first is the construction of the firm's knowledge stock, and the second is the interpretation of existing estimates of the rate of return to basic and applied research.

1.A. Lagged R&D weights in the firm's knowledge stock: When constructing a firm's research capital ( $K_t$ ) from its past R&D spending ( $R_{t-j}$ ), the usual framework (e.g. Griliches 1986) is:

$$(1) \quad K_t = \sum_{j \geq 0} \mu_j R_{t-j},$$

and the firm's output (sales, or value added) is

$$(2) \quad Q_t = [\lambda' K_t^\alpha] C_t^\beta L_t^\gamma,$$

where  $C$  and  $L$  stand for physical capital and labor. Griliches interprets  $[\lambda' K_t^\alpha]$  as follows:  $\lambda'$  is "external" technical change, and  $K_t$  is knowledge generated internally.

We derive explicit expressions for  $\lambda$  and  $K$  in terms of the parameters of the model. They both depend on how fast the firm's upgrades its technology:  $\lambda$  rises with the speed of upgrading because the adoption of better blueprints is faster. But upgrading destroys knowledge specific to abandoned technologies, accelerates the decay in the  $\mu_j$ 's, and lowers  $K$ .<sup>1</sup> To a first approximation, the  $\mu_j$ 's decay geometrically, as is indeed typically assumed in practice.

1.B. The rate of return to basic research: We distinguish applied from basic research and model an industry equilibrium in which firms that follow different research strategies coexist in the long run.<sup>2</sup> In the model the private rate of return to basic research can seem larger than that to applied research (just as Mansfield (1980), Link (1981) and Griliches (1986) found) even though it really isn't. This is because basic research signals a fast-track upgrading

---

<sup>1</sup> This point has been made before [Fellner (1970, appendix C), and Griliches (1973, fn.9)], but seems to have escaped modelling.

<sup>2</sup> Earlier work on the direction of research focused on risky versus safe research, possibly under oligopoly -- e.g. Bhattacharya and Mookherjee (1986).

strategy.

In interpreting existing estimates of the rate of return to research, we focus on Griliches's (1986) method which allows for the weight on the firm's applied research spending,  $R_t^a$ , to differ from the weight on its basic research spending,  $R_t^b$ . He defines a composite research index,  $R_t^*$ , as follows:

$$(3) \quad R_t^* \equiv R_t^a + (1 + \pi)R_t^b = R_t^*(1 + \pi s_t),$$

where

$$(4) \quad s_t \equiv \frac{R_t^b}{R_t^a + R_t^b}$$

is the share of spending on basic research in total spending on research. He interprets  $\pi$  as the premium on basic research, and his estimates of  $\pi$  (reported in his table 2) range from 0.27 to 0.52. That is, a dollar of basic research buys as much productivity as \$ 1.27 - \$ 1.52 of applied research buys, which suggests the firm might make money by shifting some of its research dollars to basic research.

## 2. The Theoretical Framework

2A. The general framework.<sup>3</sup> Let the output of the firm  $Q$ , be produced with capital,  $C$ , labor,  $L$ , and materials,  $M$ . Let  $q$  denote an efficiency parameter, so that the production function is:

$$(5) \quad Q = F(C, L, M; q) .$$

How  $q$  enters  $F(\cdot)$  will depend on the factor of production in which the learning is "embodied".<sup>4</sup> The model of the growth of  $q$  that we shall now present is, we hope, applicable to various types of learning.

Assume that  $q$  depends (a) on the technological grade  $n$  that the firm chooses to use,

---

<sup>3</sup> We have used this framework elsewhere to study optimal choice of technologies (J-N 1994), learning curves on a given technology (J-N 1995a) and the training of inexperienced workers by experienced workers (J-N 1995b).

<sup>4</sup> Bakh and Gort (1993, especially Table 4) ask how  $q$  enters the production function of the plant. If productivity growth was due to better management, we might write  $Q = qF(C, L, M)$ , as in (3). If it was due to better machines, one would write  $Q = F(qC, L, M)$ .

and (b) on how closely a grade-specific decision,  $z_n$ , matches an ideal,  $y_n$  :

$$(6) \quad q = G(y_n - z_n ; n) .$$

Suppose that  $q$  is largest when  $z_n = y_n$  . We may think of  $(y_n - z_n)$  as a mistake that reduces  $q$  below its maximal level of  $G(0, n)$ . For example  $n$  might stand for the grade of computer equipment to use, while  $z_n$  might stand for the speed of the production line in a factory, the organization of the workspace, the types of raw materials purchased, the type of labor hired and its assignment to the tasks, the choice of factor proportions, and so on. All of these aspects of operating decisions are collapsed into a scalar  $y_n$ , which fluctuates randomly over time around a mean, denoted by  $\theta_n$  .

If the firm could learn  $\theta_n$  by copying the production methods of another firm that uses grade  $n$ , or if it could copy that other firm's  $z_n$ , the firm would not need to do any research. Indeed, there are features of technologies that can be reverse engineered, and are even explained in manuals, and that do not have to be learned through costly research effort or by trial and error. But when the technology is new and is not used by other firms, the firm in question will have to discover it on its own. And even when the technology is in widespread use, (as, for example, hybrid corn was) other features of the optimal decision  $z_n$  are specific to the situation at hand, to the nature of the factors of production, the raw materials, the workspace, the specifications of the output, and so on. These, the firm will have to infer for itself. <sup>5</sup>

2B. Productivity growth on a given technological grade: Suppose that

$$(7) \quad q_n = G(y_n - z_n ; n) = \gamma^n [1 - (y_n - z_n)^2] .$$

Here  $n \geq 1$  is an integer. The maximal level of  $q_n$ , attainable under ideal conditions, is  $\gamma^n$ . Each period  $\tau$  leads to a new value of  $y_n$  :

$$(8) \quad y_{n, \tau} = \theta_n + w_{n, \tau} , \quad \tau = 0, 1, 2, \dots$$

where the  $w_{n, \tau}$  are independent, normally distributed random variables with mean zero and variance  $\sigma_w^2$ . Noise,  $w$ , is essential if learning and productivity growth on a given grade  $n$  is to be gradual.

---

<sup>5</sup> Evenson and Kislev (1975) and Mansfield *et al* (1981) find that imitation is costly, although cheaper than inventing something oneself.

Let  $t$  be the cumulative number of periods,  $\tau = 0, 1, \dots, t-1$ . The firm must choose  $z_{n,t}$  before seeing  $y_{n,t}$ . Let  $E_t(\cdot)$  denote conditional expectation. A risk-neutral firm will maximize  $E_t(q_t)$  and choose the following decision:

$$(9) \quad z_{n,t} = E_t(y_{n,t}) = E_t(\theta_n).$$

This decision is dynamically optimal as well, because the information that the firm gets does not depend on the  $z$  chosen. So, we get the reduced form

$$(10) \quad q_{n,t} = \gamma^n (1 - [\theta_n - E_t(\theta_n) + w_{n,t}]^2).$$

Now let  $\text{Var}_t(\theta_n) \equiv E_t[\theta_n - E_t(\theta_n)]^2$  be the posterior variance over  $\theta_n$ . Expected  $q$  at  $t$  then is

$$(11) \quad E_t(q_{n,t}) = \gamma^n (1 - \text{Var}_t(\theta_n) - \sigma_w^2).$$

It is increasing in  $t$ , but bounded: As  $n$  gets large,  $E_t(q) \rightarrow \gamma^n [1 - \sigma_w^2]$ . To get long-run productivity growth, the firm must therefore keep upgrading its technology, and this is the process we specify next.

2.C. The transfer of knowledge across grades of technology: There is no direct cost of switching to a different grade of technology, and no adjustment costs in  $z$ . Successive grades are linked as follows:

$$(12) \quad \theta_{n+1} = (\sqrt{\alpha}) \theta_n + \epsilon_{n+1}$$

for all  $n$ , where  $\epsilon_n$  is an i.i.d. normal variate with mean zero and variance  $\sigma_\epsilon^2$ . If the firm has not yet tried grades  $n+1, n+2, \dots$ , eq. (12) implies that

$$(13) \quad \text{Var}_t(\theta_{n+1}) = \alpha \text{Var}_t(\theta_n) + \sigma_\epsilon^2,$$

If  $\alpha = 1$  and  $\sigma_\epsilon^2 = 0$ , then  $\theta_n = \theta_k$  for all  $k$ , which corresponds to the case where knowledge is general and freely transferable across grades of the technology. But if  $\alpha = 0$ , knowledge is grade-specific.

We assume that at date  $t = 0$ , the firm finds itself on grade  $n = 1$ , and that it has never used a higher grade. We assume that the prior over  $\theta_1$  at date 1 is normal. Eq (8), and the normality of  $w_{nt}$  imply that the posterior belief at each date over the parameter of any grade,  $\theta_n$ , will also be normal. If the prior variance over  $\theta_n$  is  $x$ , then spending period  $t$  on

grade  $n$  and hence observing  $y_{n,t}$  gives, via Bayes rule, a posterior variance of  $h_1(x)$ , where

$$(14) \quad h_1(x) \equiv \sigma_w^2 x / (\sigma_w^2 + x).$$

On the other hand, if at date  $t$  the firm chooses instead to produce using the next-highest grade, grade  $n + 1$ , its prior variance over  $\theta_{n+1}$  would, in view of (13), be given by  $h_2(x)$ , where

$$(15) \quad h_2(x) \equiv \alpha x + \sigma_\epsilon^2,$$

and at the end of the period, following an observation of  $y_{n+1}$ , the posterior variance over  $\theta_{n+1}$  would then be  $h(x)$ , where

$$(16) \quad h(x) \equiv h_1(h_2(x)).$$

2. D. **Research:** We assume that the firm can direct its scientists to research any grade  $k \geq n$ . If it directs them to  $k = n$ , we shall call such research applied, whereas if it directs them to  $k > n$ , we shall call such research basic (more on this shortly). If, at  $t$ , research takes place on grade  $k$ , the firm observes a signal

$$(17) \quad y_{k,t} = \theta_k + \eta_{k,t},$$

where  $\eta_{k,t}$  is normally distributed with mean zero and variance  $\sigma_r^2$ . Three points should be noted about how we are modelling research:

- (i) Information from research is the same as the information obtained under "learning by doing" except that the precision of the information from research,  $1/\sigma_r^2$ , may differ from the precision obtained from learning by doing,  $1/\sigma_w^2$ .
- (ii) Research dollars spent this period yield information that one can use next period. Given the lags involved in the maturation of research projects (see Pakes and Schankerman 1984), this means that the length of the period should be fairly long -- at least a year.
- (iii) Unless  $\alpha = 0$ , this signal will provide information on the parameter  $\theta_n$  of all grades of technology, and not only on grade  $k$ . This includes grades  $n$  that are strictly less than  $k$ .

### 3. Basic and Applied Research, and the Choice of an Upgrading Policy

We define applied research to be spending aimed at learning more about the technological process a firm is already using, or about a good that it is already producing.

Basic research, on the other hand, is directed towards processes not yet in use, or goods not yet produced. Our definition of basic research agrees with that of the National Science Foundation (1959, p. 124) -- "...original investigation for the advancement of scientific knowledge....which do(es) not have immediate commercial objectives." It is also consistent with some findings in Mansfield's (1981) study of the effects of basic research among chemical and petroleum firms. In eq.(2) Mansfield reports the finding that holding a firm's R&D expenditures constant, its innovative output is positively and significantly related to the share of its R&D spending going to basic research.

After (a) making some basic assumptions, this section will (b) establish a direct relation between the firm's research policy and its upgrading policy, and (c) characterize the optimal upgrading policy.

3.A. A preliminary assumption about upgrading: We begin by assuming that the firm will upgrade at a constant rate (in section 4 we shall turn this assumption into a result). Fix a date  $t$ . Suppose that a firm currently operates grade  $N$  of technology -- the highest it has ever operated. Assume that the firm upgrades its technology at the constant rate  $u$ , so that, since the initial grade has been normalized to be zero,

$$(18) \quad N = u t.$$

The firm's long-run growth in  $q$  -- call it  $g$  -- is:

$$(19) \quad g = \gamma^u - 1 \approx u \ln(\gamma).$$

If we knew  $\gamma$ , we could identify  $u$  from the firm's long-run rate of Productivity growth:

$$(20) \quad u \approx \frac{g}{\ln(\gamma)}.$$

Suppose then that we identify basic research as associated with a "high  $u$ " upgrading strategy, and applied research with a "low  $u$ " upgrading strategy. More precisely research ahead of the firm's current technological position is basic, and research on refining the firm's current technology is applied.<sup>6</sup> This will allow us to explain the puzzling finding by Mansfield

---

<sup>6</sup> Evidence supporting this view is in Mansfield (1981), who finds that innovative firms put more weight in their research portfolios on basic research.



(1980), Link (1981), and Griliches (1986) that basic research higher private rate of return than applied research. In our model, given (20) and our definition of basic research, the productivity growth of firms doing basic research will for ever exceed that of firms doing applied research.

3.B. The relation between the research policy and the upgrading policy: To simplify the argument, suppose that the firm maximizes the expected return in the current production period, and that it directs its research with this sole aim in mind. In other words, assume that the firm's discount factor is zero (the result easily generalizes to the case in which the discount factor is positive, but small). Proposition A.1 in the appendix shows that under mild restrictions on the parameters, all of the firm's research effort would optimally be directed towards that technological grade which it is about to produce output with. Under these restrictions, the upgrading policy and the research policy are one and the same.<sup>7</sup>

The question now is: Why would two firms choose such different research strategies with the resulting permanent difference in their productivity growth? We address this next.

3.C. The optimal upgrading policy: Fix a date and let  $x$  be the before research posterior variance on the status quo grade  $N$ . Define  $h_r(x) = \sigma_r^2 x / (\sigma_r^2 + x)$ , the learning by research function. We shall normalize the payoffs by dividing them by  $\gamma^N$ . Section 3B says that the firm will research the same grade it is just about to produce on. If the firm sticks to its status quo technology, its research will yield a posterior variance  $h_r(x)$ , and its expected return will be

$$\ell(x) \equiv 1 - \sigma_w^2 - h_r(x).$$

Assume that an upgrade can only be made by exactly one step. If it switches to the next grade and researches it before producing, the firm's posterior variance on that grade will be  $h_r[h_2(x)]$ , and its expected return will be

$$r(x) \equiv \gamma(1 - \sigma_w^2 - h_r[h_2(x)]).$$

The functions  $\ell(x)$  and  $r(x)$  are downward-sloping and convex in  $x$ . In particular, fig. 2 assumes that there is a value  $\hat{x}$  such that the following are true on  $[0, \hat{x}]$ :  $r(0) < \ell(0)$ ,  $\ell(x)$

---

<sup>7</sup> Of course, this statement can not literally be true in that the typical research project takes a year or two to bear fruit in the form of sales (Pakes and Schankerman, 1984). The statement can make sense only if the period is long enough.

curve is steeper than the  $r(x)$  curve and  $r(\hat{x}) > \ell(\hat{x})$ . (In the appendix we explain the conditions on the parameters which results in this.) There will therefore exist an  $x^*$  in  $(0, \hat{x})$  such that  $r(x) < \ell(x)$  for  $x$  on  $[0, x^*)$  and  $r(x) > \ell(x)$  for  $x$  in  $(x^*, \hat{x}]$ , as shown in figure 1.<sup>8</sup>

Suppose that a firm has posterior variance of  $x$  on the status quo grade. Suppose that the firm researches and then produces on the next grade. Its posterior variance on the new

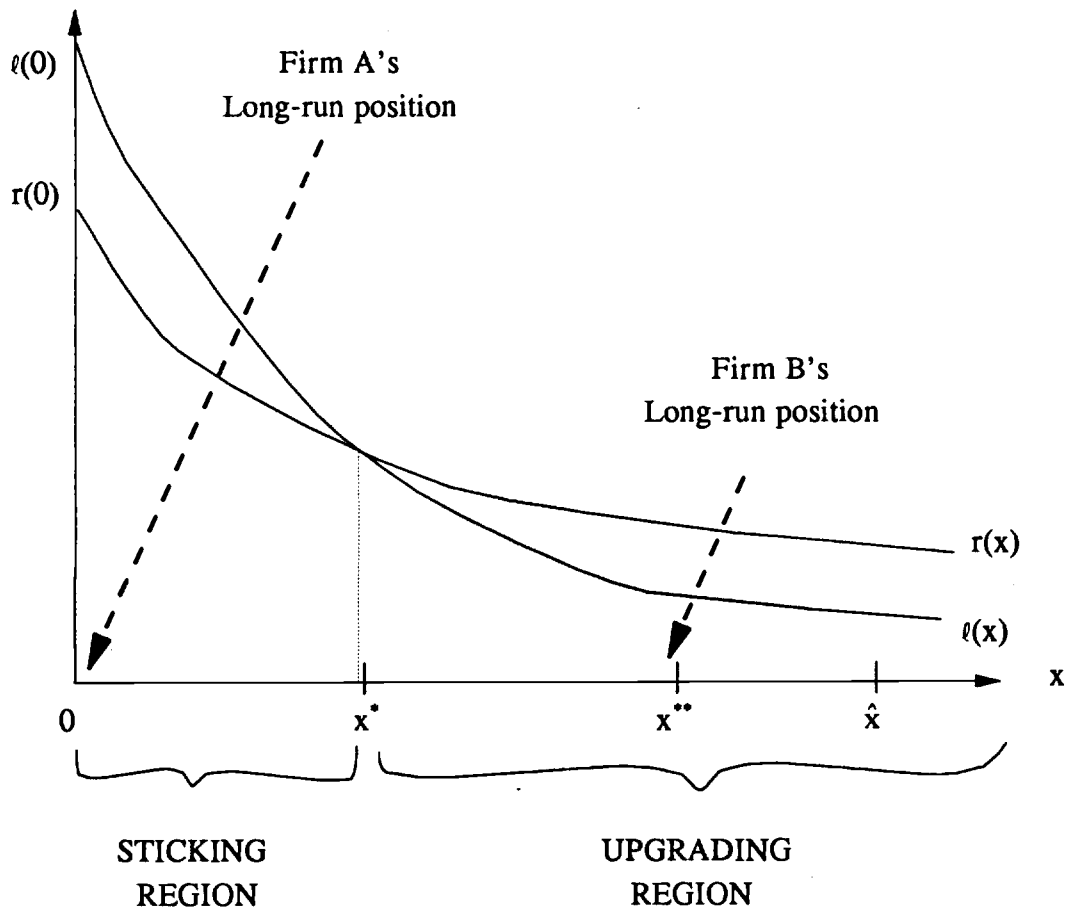


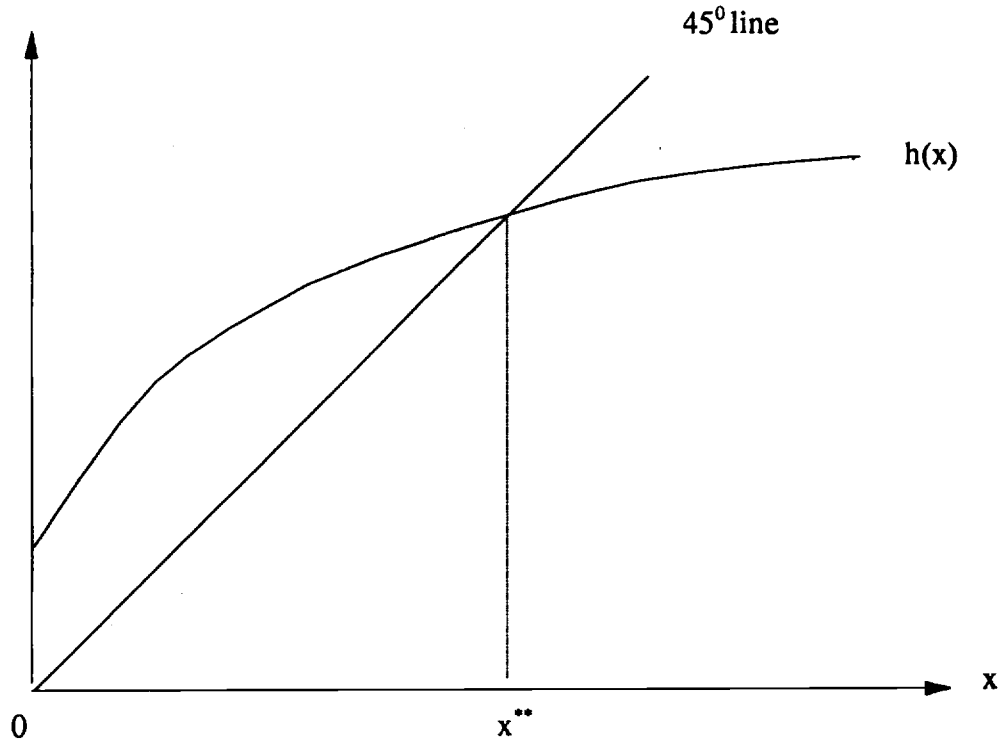
Figure 1

<sup>8</sup> Figure 2 shows that the set  $[0, \hat{x}]$  is stable and absorbing: From any initial value, the posterior variance will enter  $[0, \hat{x}]$  in finite time (regardless of the policy that is chosen) and once in that set, it remains there.

grade will therefore be, after observing the learning by doing signal, given by the function

$$h(x) \equiv h_1(h_r(h_2(x))).$$

The  $h$  function can be shown to be strictly concave with  $h(0) > 0$  and  $h(\infty) < \infty$ , as drawn in Figure 2. It therefore has a unique positive fixed point which we denote by  $x^{**}$ .



**Figure 2:** The evolution of  $x$  for an upgrading firm

Implicit in figure 1 is our assumption that  $x^* < x^{**}$ . If the firm upgrades in each period the dynamics of the posterior variance will be given by iterates of the  $h$  function, which will converge monotonically to  $x^{**}$ . Fig. 2 shows the following: When  $x$  is low, and in particular when  $x \leq x^*$ , the firm chooses to stick in each and every period. This will cause the posterior variance to converge to zero. When  $x$  is large, (i.e.,  $x \geq x^*$ , the firm will choose to upgrade in every period. Figure 2 shows that the posterior variance then converges to  $x^{**}$ .

We summarize these results of section 3.C. in the following proposition:

**Proposition 3.C.1.** Assume the payoffs are as in fig.2. Suppose two firms, I and II, are using the same grade  $n$  and have initial posterior variance,  $x_0'$  and  $x_0''$  respectively, with  $x_0' < x^* < x_0''$ . Then firm I initially has a higher expected output, but it will never upgrade. Firm II, will keep upgrading, and growing forever.

These ideas will be central to the model in the next section.

#### 4. A long-run, "coexistence" equilibrium.

We now look at equilibrium. Since the productivity of the upgrading firm keeps growing relative to that of the other firm, two related questions arise: (A) wouldn't the laggard eventually be driven out? and (B) would we not therefore have only upgrading firms in an equilibrium, and therefore no heterogeneity in the long run in the type of research being done? We now model an equilibrium in which the answer to (A) is yes, but the answer to (B) is no.

Suppose that the economy grows as  $\gamma^t$ . The demand for the industry's product therefore grows as  $\gamma^t$ . We shall now show that there is a steady state equilibrium in which two kinds of firms coexist, in which industry price and the number of firms of each type is constant, and in which the industry output supplied also grows as  $\gamma^t$ .

At date  $t$ , the aggregate quantity  $Q$  is related to the industry's price  $p$  by the demand curve  $Q = \gamma^t D(p)$ . At  $t = 0$ , suppose the status quo for all firms is  $n = 0$ . As far as  $x$  is concerned, firms are of 2 types: A, and B; the initial beliefs of type A's are  $x = 0$ , while those of type B's are  $x = x^{**}$ . Assume that the situation is as depicted in figure 1. Then type A's always stick to their old technology, hence their  $x = 0$ , always. Type B's always upgrade, hence their  $x = x^{**}$ , always. We shall look for a steady state in which firms are (essentially) of 2 types, A (the "transients") and B (the upgraders), having measures  $\mu_A$  and  $\mu_B$ . If the product price can be shown to be constant over time, industry demand will grow as  $\gamma^t$ , and the population of upgrading firms will maintain a fixed share of that demand,  $\mu_B/(\mu_A + \mu_B)$ .

Entry and exit of transients: Entry costs at date  $t$  are  $\gamma^t c_0$ , and operating costs are  $\gamma^t c_1$ . At date  $t$ , we assume that there is an infinite number of potential entrants who understand perfectly well the previous period technology used as the status quo of the industry leaders.<sup>9</sup> This grade is grade  $n = t - 1$ , and for these potential entrants,  $x = 0$  on this grade. For simplicity, assume that the discount factor is zero. Since we are in a situation as

---

<sup>9</sup> This simplifies the analysis a lot. In general, what we need here is that the entrants have somehow managed to learn the previous period technology well enough to put them in the "stick" region of figure 1.

depicted in figure 1, these potential entrants will use grade  $t - 1$  if they decide to enter. Let  $p_t$  for the expected returns to entry to be zero at date  $t$ , we need that

$$(21) \quad p_t \gamma^{t-1} (1 - \sigma_w^2) = \gamma^t (c_0 + c_1),$$

which means that

$$(22) \quad p = \frac{\gamma(c_0 + c_1)}{1 - \sigma_w^2}.$$

A transient firm exits as soon as its operating costs,  $\gamma^{t+T} c_1$ , exceed  $p \gamma^t (1 - \sigma_w^2)$  -- its expected revenue on the technological grade it chose when it entered. That is, its lifetime in the industry is the smallest integer  $T$  satisfying:

$$(23) \quad p(1 - \sigma_w^2) \leq \gamma^T c_1.$$

In the equilibrium that we shall construct, the measure of upgraders is  $\mu_B$ , and the measure of transients is  $\mu_A$ . However, there are different types of transients, one for each age (0, 1, ..., T-1), and the measure of each different type is  $1/T$ . Therefore, normalized by  $\gamma^t$ , steady state output demanded must equal output supplied:

$$(24) \quad D(p) = \mu_A \left\{ T^{-1} \sum_{t=1}^T \gamma^{-t} \right\} (1 - \sigma_w^2) + \mu_B (1 - \sigma_w^2 - x^{**}).$$

The term in braces multiplied by  $(1 - \sigma_w^2)$  is the average productivity of transients. The measure  $\mu_B$  is a parameter, given by the initial conditions -- the initial number of type B firms.

An equilibrium is a vector  $(\mu_A, \mu_B, p, \text{ and } T)$  of nonnegative scalars satisfying (22), (23) and (24). A further requirement that equilibrium must satisfy is that the type B firms do not want to exit:

$$(25) \quad p(1 - \sigma_w^2 - x^{**}) > c_1.$$

There is a continuum of coexistence equilibria here, and it helps to think of them as being indexed by  $\mu_B$ . We can imagine that at date zero, there is a measure of  $\mu_B$  type-B

firms in the industry. Thus  $\mu_B$  can be thought of as an initial condition. Since  $T$  and  $p$  do not depend on either  $\mu_A$  or  $\mu_B$ , there is a linear, negative relation between  $\mu_A$  and  $\mu_B$  as we go across equilibria. This linear relation emerges from the condition (24) that defines  $\mu_A$  implicitly in terms of  $\mu_B$ . Note that the coexistence equilibrium involves a steady entry and exit of transients in each period.

But if the initial condition  $\mu_B$  is large, equilibrium will not be of the coexistence type, and we will have  $\mu_A = 0$  -- there are so many upgraders that there is no room in the industry for any transients. In this case,  $p$  will no longer be given by (22), because type A's will now strictly prefer not to enter. Instead, it is the upgraders who will be indifferent between staying in and exiting, (25) will hold as an equality, and  $\mu_B$  will have a unique solution:

$$(26) \quad \frac{1}{(1 - \sigma_w^2 - x^{**})} \quad D\left(\frac{c_1}{(1 - \sigma_w^2 - x^{**})}\right)$$

So, there is a unique “non-coexistence” equilibrium in which there are no transients, and hence this steady state equilibrium involves no entry and exit.

This section has shown how that there can exist an equilibrium in which two types of firm can coexist. In any such equilibrium, firms that opt for basic research and who also opt for a fast-upgrading, high growth strategy, and at the same time, there are transient firms that choose to cash in on their grade-specific knowhow.

## 5. The observed relation between productivity, and basic and applied research

5.A. Interpreting productivity growth regressions: In the above equilibrium, basic research is positively correlated with the speed of upgrading and this can cause a bias on the coefficient of basic research in an OLS productivity regression. In fact, combined with the argument of section 3.B, this definition of basic research implies

**Proposition 5.A.1.** Suppose that the conditions of Proposition A.1 (in the appendix) are met, and that the firm is interested in maximizing its current payoff. A firm that always does basic research will in the long run grow faster than a firm that always does applied research.

Let's see how this bears on the actual panel data regressions. Suppose a sample of firms is drawn from the long run equilibrium modelled in section 4. From (19) we find that since its  $u_B = 1$ , a type B firm's long-run growth rate is  $\ln(\gamma)$ , while a type A firm's long-run productivity growth is zero, because  $u_A = 0$ . That is,

$$(27) \quad g_{i,t} = [\ln(\gamma)] u_{i,t}$$

where each firm  $i$  is of type either A or B, and  $u_{i,t}$  is its basic research intensity at date  $t$ . In the steady state equilibrium that section 3.C. presents,  $u_{i,t} = 0$  for all  $t$  if  $i$  is of type B, and  $u_{i,t} = 1$  for all  $t$  if it is of type A. Equation (27) corresponds to the productivity growth regressions reported in table 3 of Griliches (1986), with the variable  $u$  corresponding to his variable "BR 72", and to what we call  $s_{i,t}$  in (4). This variable comes in with a coefficient between 0.034 and 0.056. Since the dependent variable is measured in percentage points of growth per year, this is a very large effect.

The argument can also be put in terms of the "omitted variables bias" language. Basic research is positively correlated (in our example perfectly correlated) with the speed of upgrading, and this is an omitted endogenous regressor. Hence we estimate a positive  $v_3$ . Indeed, Griliches (1986, p. 153) considers this to be one possible interpretation of his results when he says that firms have different opportunities to do basic research. In our paper, research opportunities are determined by initial conditions -- a firm's initial knowledge effectively changes its opportunities to doing applied and basic research. So our model provides a context in which this interpretation is exactly right.

5.B. Other possible explanations: Other possible explanations for the basic research coefficient have been mentioned, and we list them here.

5.B.i Griliches (1986) mentions two other reasons why the returns to basic research would appear high: (A) Basic research is probably riskier than applied research, and so the higher coefficient could be a risk premium to basic research, and (B) the higher rate of return to basic research could represent a higher social but not a higher private return -- this could occur if the basic research spenders learn from other basic research spenders.

5.B.ii Nelson (1988) provides a "coexistence" model with firms that follow different research strategies -- innovation and imitation. An imitator acquires an innovator's invention after a  $\tau$  - period delay. It is therefore natural in Nelson's model to think of innovators as doing basic research, and of imitators as doing applied research. All firms grow at the same rate; the productivity differential between the two types of firm is in levels only. Therefore, while the Nelson model can help us understand the results of table 2 of Griliches (1986) which are about levels of productivity, it is not consistent with the results in Griliches's table 3 that show a strong association between the share of basic research and productivity growth. The latter finding is confirmed in table 3 of the study by Lichtenberg and Siegel (1991).

5.C.iii There could be reverse causation: More productive firms may do more basic research as a "luxury item". Certainly, productive firms generate larger cash flows, and there is some evidence in Hall (1990) and elsewhere that cash flow seems to affect R&D budgets. As with Nelson's model, however, this argument can shed light on the results of table 2 of Griliches (1986), but not, it would seem, those in his table 3 that show a positive relation between the share of basic research in a firm's R&D budget, and its productivity growth.

5.C. Auxiliary implications: Our interpretation is supported by Mansfield's (1981) finding that innovative firms skew their research portfolios towards basic research. But there are other, yet untested implications: Technology is usually embodied in equipment. Hence upgrading should be accompanied by equipment investment. A test of proposition 5.A.1 would then check if firms doing more basic research also engage in more equipment investment. Tangentially related to this is Lach and Schankerman's (1989) finding that in the time series at the firm level, total R&D raises total investment. And Miyazaki (1994, esp. table 7), using data on company research disaggregated by type, finds a lot of persistence, among close product market competitors, in the type of research carried out. This indirectly supports the implication that firms that start out doing applied research will always do applied research, and similarly for firms doing basic research.

The model of section 4 is consistent with the finding by Doms *et al* (1994) that firms using advanced technologies have higher survival rates. The exit hazard is zero, except at age T, where it is one. There is some evidence of a non-monotone hazard of plant exit in Table 5 of Agarwal and Gort (1993).

## 6. The Index of Research Capital

The  $\mu$ 's are hard to estimate, because a firm's R&D spending tends to be smooth over time (Hall 1995). Most researchers therefore assume geometrically declining  $\mu$ 's -- typically at 15%. Some have inferred obsolescence from the decay in the number of citations that a patent gets over time, or the number of citations of academic articles.<sup>10</sup> And others have tried to infer depreciation and obsolescence by following the prices of used equipment of various vintages.<sup>11</sup> We now derive the  $\mu_j$  in our model.

Fix a date  $t$ . Suppose that a firm is currently operating a grade  $N$  of technology.

---

<sup>10</sup> Adams (1990), Caballero and Jaffe (1993), Trajtenberg, Henderson and Jaffe (1992).

<sup>11</sup> Hulten and Wykoff (1981), Oliner (1993) are interested in depreciation of equipment, but in new technologies are embodied in equipment, then their estimates say something about the depreciation of research dollars too, so long as one recognizes the time that the research dollar has to "wait" before its fruits are embodied in a machine.



Suppose that as of date  $t$ , for  $m = 1, \dots, N$ , there have been  $\tau_m$  units of learning by doing signals (i.e., the firm operated grade  $m$  for  $\tau_m$  periods), and there have been  $\rho_m$  units of research on grade  $m$ . What is the number,  $\rho_m'$ , of learning by doing signals that this is equivalent to? Well, we require that for each initial posterior variance  $x$ ,

$$(28) \quad 1/x + \rho_m'/\sigma_w^2 = 1/x + \rho_m/\sigma_r^2 .$$

Then  $\rho_m' = [\sigma_w^2/\sigma_r^2]\rho_m$ . Hence a research signal is equivalent to  $\sigma_w^2/\sigma_r^2$  learning by doing signals. Seeing  $\tau_m$  learning by doing signals and  $\rho_m$  research signals is the same as seeing  $R_m$  learning by doing signals, where

$$(29) \quad R_m = \tau_m + [\sigma_w^2/\sigma_r^2] \rho_m .$$

We shall call  $R_m$  effective signals.

6.A. The relation between  $R_m$  and  $R_t^*$ : As defined in (29),  $R$  corresponds closely to  $R^*$  in (28), but for three differences. First, the units in which they are measured are different:  $R$  is in units of numbers of signals, while  $R^*$  is in units of real dollars. This is not a problem if the price of research relative to other goods is constant, or if we have the correct price index with which to deflate gross R&D investment at each date.

Second,  $R_t^*$  has a time subscript, whereas  $R_m$  has a grade subscript. Hence to relate the two, we need to know what grade was researched at what date. This in turn will be related to the speed with which the firm upgrades its technology.

Third,  $R_m$  contains signals  $\tau_m$  that the firm did not purchase through its research outlays. How much information the firm acquires in this way again depends on the its upgrading policy: A firm that upgrades rapidly will have fewer such signals per grade because it will be spending less time on any one grade.

If we can somehow identify the speed with which the firm upgrades its technology, and if we can adjust for these measurement problems, our theory of how the firm's knowledge relates to its effective signals ( $R_m$ ) will have implications for how the firm's research capital stock depends on its past R&D spending, ( $R_t^*$ ). This is the task to which we now turn.

6.B. R&D signals on past grades only: In general, the firm will have signals on grades that it is not currently using. But we shall begin with the simpler case in which such signals are not available. Then we shall generalize the index to allow for the additional information about future grades.

Let  $N$  denote the grade on which production is about to take place. Suppose that at

date  $t$  there have been  $R_m$  effective signals on grade  $m$  for  $m = 1, \dots, N$  [with each  $R_m$  coming from any combination of learning by doing and research satisfying (29)]. Let us suppose that there has been no information on any grade in excess of  $N$ . Suppose that before any information was received, the initial prior variance over  $\theta_1$  at date 1 is  $x_1$ . The expected productivity on grade  $N$  will then be

$$(30) \quad Eq_N = \gamma^N (1 - \sigma_w^2 - x'_N),$$

where<sup>12</sup>

$$(31) \quad x'_N = h_1^{R_N} (h_2 [h_1^{R_{N-1}} (\dots (h_1^{R_2} (h_2 [h_1^{R_1} (x_1)]) \dots))]).$$

This is the general expression for the stock of knowledge about  $\theta_N$  when R&D signals are not available on any grade higher than the currently used grade  $N$ .

How do (30) and (31) relate to (1) and (2)? The correspondence would be quite close if we could define

$$(32) \quad \lambda^t K_t = E_t(q_N).$$

The presence of the trend  $\lambda^t$  means not only that there is upgrading, but that it occurs regularly. Earlier we assumed upgrading at a constant pace:

$$(33) \quad N = ut,$$

Then (32) holds with the following definitions for  $\lambda$  and  $K$ :

$$(34) \quad \lambda = \gamma^u, \quad \text{and} \quad K_t = 1 - \sigma_w^2 - x'_N.$$

Our model therefore gives a specific meaning to the production function in (1). Does (2) also have an interpretation? The answer is yes, in terms of a linear approximation of  $K_t$ :

---

<sup>12</sup> To handle the possibility that  $R$  is zero for some grades that were never used or researched, we adopt the convention that  $h_1^0(x) = h_t^0(x) = x$ .

**Theorem 6.B.1** (The Linearized Research Capital Index  $K^*$ ): Fix an  $R$  and an initial variance  $x_1$ .<sup>13</sup> Then under the definition of  $K_t$  in (34),  $K_t$  can be linearly approximated by  $K_t^*$ , where

$$(35) \quad K_t^* = A \sum_{m=1}^N \delta^{n-m} R_m + c = A \sigma_w^2 \sum_{m=1}^N \delta^{n-m} (a \tau_m + b \rho_m) + c$$

is the Taylor series linear expansion of (34) around  $R_m = R \forall m$  when  $x_1$  is close to  $\bar{x}$  [the fixed point of the map  $h_2(h_1^R(\cdot))$ ]. On the right-hand side of (35), the parameters  $a$ ,  $b$ ,  $c$ , and  $A$  are given by

$$(36) \quad a = 1/\sigma_w^2; \quad b = 1/\sigma_r^2; \quad c = 1 - \sigma_w^2 - \bar{x} - AR \sum_{m=1}^N \delta^{n-m}, \quad A = [h_1^r(\bar{x})]^2 / \sigma_w^2,$$

and the discount factor  $\delta$  is strictly less than one and is given by

$$(37) \quad \delta = \alpha h_1^R(\bar{x})^2 / \bar{x}^2.$$

**Proof:** In the appendix. ■

From the definition of  $h_1$  in (14), one can show that  $h_1^R(\cdot)$  is an increasing and concave function, and so the ratio  $h_1^R(\bar{x})/\bar{x}$  is decreasing in  $\bar{x}$ . Since  $\bar{x}$  depends on the upgrading policy, so will  $\delta$ . We shall analyze some properties of  $K_t^*$  in section 4. Next, we show that the index retains its basic character even when the firm has signals over grades that it has not yet used.

**6.C. Signals on past and future grades:** Now suppose that in addition to signals on grades  $n$  used in the past there are also signals about grades that have not yet been used. Let  $x_N'$  denote, as before, the posterior variance on the current grade, grade  $N$ , given only signals on past grades. Eq. (31) and Theorem 6.B.1 above are expressions for  $x_N'$ . Suppose that  $\rho_{N+m}$  units of effective signals are obtained on grade  $N+m$  for  $m=1,2,\dots$ . To make headway we shall suppose that all the future signals are concentrated on one grade, grade  $N+k$ , and that there are  $\rho_{N+k}$  such signals.

**Theorem 6.C.2** (The Linearized Research Capital Index  $K^{**}$ ): Let  $R$ ,  $x_1$  be as in Theorem 6.B.1. Then  $K_t$  can be linearly approximated by  $K_t^{**}$ , where

---

<sup>13</sup> Implicit in the following formula is fact that  $\alpha \neq 0$ .

$$(38) \quad K_t^{**} = c_1 K_t^* + c_2 \rho_{N+k} + c_3,$$

where the approximation is around the points stated in Theorem 6.B..1 and around  $\rho_{N+k}$  close to some  $\bar{\rho}$ , and where  $c_1 = \partial G/\partial x$ ,  $c_2 = -\partial G/\partial \rho$  (both evaluated at  $(x, \rho) = (\bar{x}, \bar{\rho})$ ), and  $c_3 = (1-c_1)(1 - \sigma_w^2) - G(\bar{x}, \bar{\rho}) + c_1 \bar{x} - c_2 \bar{\rho}$ .

**Proof:** In the appendix. ■

So, the geometric discounting of past research signals is present in  $K_t^{**}$  just as it was in  $K_t^*$ .

## 7. Endogenous obsolescence: Upgrading and the research capital index

The firm's upgrading speed  $u$  will raise the depreciation rate by lowering  $\delta$ . Before showing this, we first establish that changing  $u$  is equivalent to keeping  $u$  fixed and changing some of the other parameters.

7.A. An invariance property: Redefining the size of the step turns out to be equivalent to changing some other parameters:

**Proposition 7.A.1:** (re-scaling of the technological grid) Suppose firm A upgrades by one step each period, while firm B upgrades by  $u$  steps. Suppose they both start on grade  $n = 1$ , and with the same  $x_1$ . If

- i) Firm B faces the parameters  $\gamma_B$ ,  $\alpha_B$ , and  $\sigma_{B\epsilon}^2$ , and effective signals  $R_{B, um}$  (for  $m = 1, 2, \dots$ )
- ii) Firm A has the parameters  $\gamma_A = \gamma_B^u$ ,  $\alpha_A = \alpha_B^u$ ,  $\sigma_{A\epsilon}^2 = (1-\alpha_B^u)/(1-\alpha_B)\sigma_{B\epsilon}^2$ , and effective signals  $R_{A, m}$  (for  $m = 1, 2, \dots$ )
- iii)  $\sigma_w^2$  is the same for both firms,

then if  $R_{A, m} = R_{B, um}$  (for  $m = 1, 2, \dots$ ) firm A's and firm B's productivities, as well as their indexes  $K^*$  and  $K^{**}$  will be the same at each date.

**Proof:** In the appendix.

7.B. Faster upgrading speeds up the depreciation of knowledge: To get the expressions in terms of  $u$  one merely replaces  $\alpha$  with  $\alpha^u$  everywhere and  $\sigma_\epsilon^2$  to  $(1-\alpha^u)/(1-\alpha)\sigma_\epsilon^2$ .

**Proposition 7.B.2:** An increase in  $u$

- i) Lowers  $K$  [in (34)] by raising  $x'_N$  [in (31)] for any  $x_1 < \sigma_e^2/(1 - \alpha) \equiv \hat{x}$ ,
- ii) Leaves  $a$  and  $b$  unaffected [in (36)],
- iii) Lowers  $\delta$  [in (37)] -- the relative weight of one period lagged versus current signals.

**Proof:** In the appendix

The faster depreciation of knowledge and the reduction in  $K$  is the result of a faster depreciation of technologically specific knowhow when the step-size  $u$  is larger. This is the downside of rapid technological upgrading -- we never stick with anything long enough to understand it perfectly, and when we upgrade to a relatively advanced method, we find that we do not understand it too well. The upside of a larger  $u$  is that our log-run productivity growth is higher.

## 8. Conclusion

In this paper, we have postulated an explicit formulation of how information enters the production function, and we have used it to answer two questions. The first question concerns the distinction between basic and applied research, and how differences in firms' research portfolios might in be related to their productivities. We have shown how different research portfolios can be an optimal policy for similarly situated firms, and that firms that follow different policies can coexist in a long-run equilibrium. We have shown how basic research can seem to have a higher rate of return than applied research. The reason is that firms doing basic research are upgrading their technologies at a faster rate than firms doing applied research.

The second question we have looked at is the construction of the firms knowledge stock. We find that as a first approximation, one can assume (as is typically done in practice) that the lags decay geometrically. The decay rate -- the rate of obsolescence -- is not exogenous, however. It depends on how fast new technological grades are coming into use. Qualitatively, this point is well known, of course; we have simply derived an explicit formula for how obsolescence works.

## References:

- Adams, James, "Fundamental Stocks of Knowledge and Productivity Growth," Journal of Political Economy 98 (1990): 673 - 702.
- Agarwal, Rajshree, and Michael Gort, "The Evolution of Markets and Entry, Exit, and Survival of Firms," Discussion paper 9307, SUNY Buffalo, August 1993.
- Bakh, Byong-Hyong and Michael Gort. "Decomposing Learning by Doing in New Plants." Journal of Political Economy 101, no. 4 (August 1993): 561 - 583.
- Bhattacharya, Sudipto, and Dilip Mookherjee, "Portfolio Choice in Research and Development," The Rand Journal of Economics 17, no. 4 (Winter 1986): 594 - 605.
- Caballero, Ricardo, and Adam Jaffe, "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," NBER Macroeconomic Annual, 1993.
- Chari, V.V., and Hugo Hopenhayn, "Vintage Human Capital," Journal of Political Economy 99, no. 6 (December 1991): 1142 - 65.
- Chow, Gregory, Econometrics, New York: McGraw-Hill, 1983.
- Doms, Mark, Timothy Dunne, and Mark Roberts, "The Role of Technology Use in the Survival and Growth of Manufacturing Plants," U.S. Bureau of the Census, Center for Economic Studies, 1994.
- Evenson, Robert, and Yoav Kislev, Agricultural Research and Productivity New Haven: Yale University Press, 1975
- Fellner, William, "Trends in the Activity Generating Technical Progress," American Economic Review 60, no. 1 (March 1970): 1 - 29.
- Foster, Andrew, and Mark Rosensweig. 1994. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." Unpublished paper, University of Pennsylvania.
- Griliches, Zvi, "Research Expenditures and Growth Accounting," in Science and Technology in Economic Growth, B.R. Williams, ed., MacMillan Press, 1973.
- Griliches, Zvi, "Productivity, R&D, and Basic Research at the Firm Level in the 1970's," American Economic Review 76, no. 1 (March 1986): 141 - 54.
- Hall, Bronwyn, "The Impact of Corporate Restructuring on Industrial R&D," Brookings Papers on Economic Activity: Microeconomics (1990): 85 - 124.
- Hall, Bronwyn, "The Private and Social Returns to R&D: What Have We Learned?" unpublished paper, UC Berkeley, April 1995.
- Hulten, Charles, and Frank Wyckoff, "The Estimation of Economic Depreciation Using Vintage Asset Prices," Journal of Econometrics 15 (April 1981): 367 - 96.
- Jovanovic, Boyan, and Yaw Nyarko. 1994. "The Bayesian Foundations of Learning by

- Doing," New York University. NBER working paper no. 4739.(J-N 1994)
- Jovanovic, Boyan, and Yaw Nyarko, "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," Brookings Papers on Economic Activity (Microeconomics Issue no.1), 1995.(J-N 1995a)
- Jovanovic, Boyan, and Yaw Nyarko, "The Transfer of Human Capital," Journal of Economic Dynamics and Control 19 (June 1995):1033 - 1064 (J-N 1995b).
- Lach, Saul, and Mark Schankerman, "Dynamics of R&D and Investment in the Scientific Sector," Journal of Political Economy 97, no. 4 (August 1989): 880 - 904.
- Lichtenberg, Frank, and Donald Siegel, "The Impact of R&D Investment on Productivity -- New Evidence Using Linked R&D-LRD Data," Economic Inquiry 29 (April 1991): 203 - 228.
- Link, Albert, "Basic Research and Productivity Increase in Manufacturing: Additional Evidence," American Economic Review 71, no. 5 (December 1981): 1111 -2.
- Mansfield, Edwin, "Basic Research and Productivity Increase in Manufacturing," American Economic Review 70, no. 5 (December 1980): 863 - 73.
- Mansfield, Edwin, "Composition of R&D Expenditures: Relationship to Size of Firm, Concentration, and Innovative Output," Review of Economics and Statistics (1981): 610 - 615.
- Miyazaki, Kumiko, "Search, Learning and Accumulation of Technological Competences; The Case of Optoelectronics," Industrial and Corporate Change 3, no. 3 (1994): 631 - 654.
- National Science Foundation, Methodology of Statistics on Research and Development, Washington, D.C., 1959.
- Pakes, Ariel, and Mark Schankerman, "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources," in R&D, Patents, and Productivity, Zvi Griliches, ed., University of Chicago Press, for NBER, 1984.
- Oliner, Stephen, "New Evidence on the Retirement and Depreciation of Machine Tools," Working paper no. 147, Board of Governors, December 1993.
- Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe, "Ivory Tower Versus Corporate Lab: An Empirical Study of Basic Research and Appropriability," NBER Working Paper no. 4146, August 1992.

## Appendix:

First, we prove the claim made in section 3B, that a myopic firm will always research that grade which it is about to use to produce with. Surprisingly, under some conditions, the firm could extract maximal information about a technological grade by researching a different grade! Further assumptions are needed to rule out this paradoxical case. They are provided in the following Proposition:

**Proposition A.1. (Monotonicity of research grades):** Fix any date  $t$ . Suppose that the status quo grade (i.e., the highest grade that the firm has ever operated) is  $N$ . Fix any grade  $n \geq N$ . Let  $\text{Var} [\theta_n | y_k]$  denote the posterior variance on grade  $n$  conditional on researching grade  $k$ . Then under condition (M) below,  $\text{Var} [\theta_n | y_k]$  is non-increasing in  $k$  for all  $k \leq n$  and is strictly increasing in  $k$  for all  $k \geq n$ , and hence attains a minimum at  $k = n$ . In particular, condition (M) guarantees that to minimize  $x_n$  the firm will choose to research that grade  $k = n$ , and not any other grade.

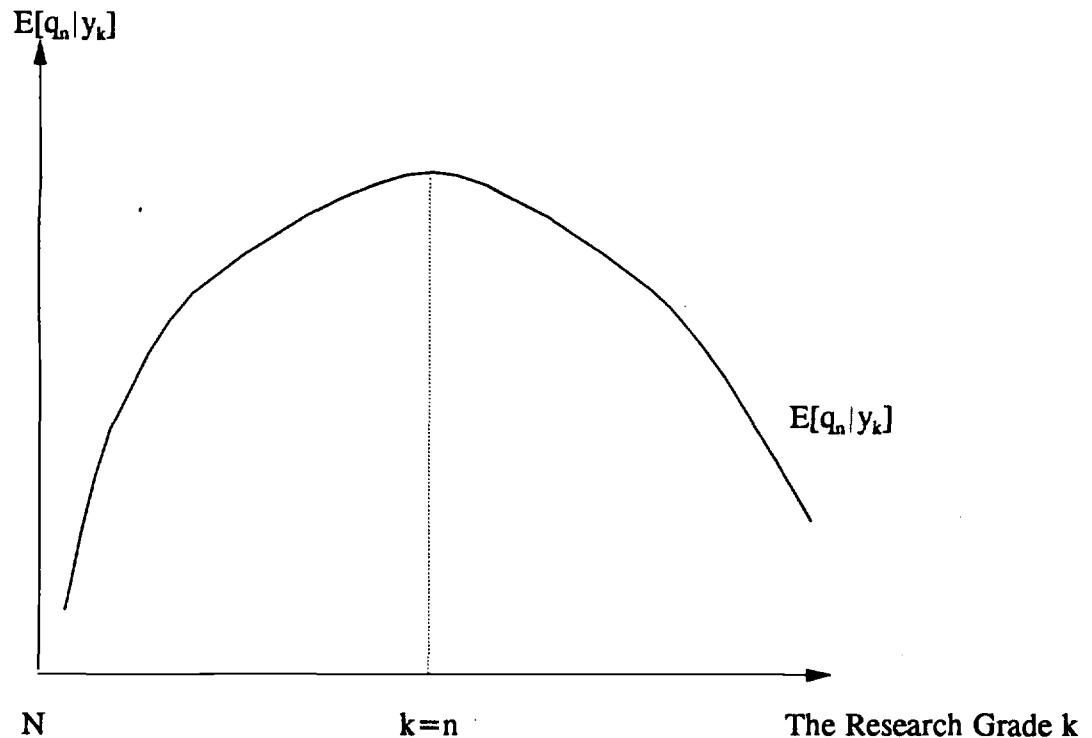


Figure A.1: the productivity,  $E[q_n | y_k]$  as a function of the research grade  $k$



- (M) **Either** (i)  $\alpha < 1$ ,  $\sigma_r^2 - \sigma_\epsilon^2/(\alpha-1) < 0$  and  $x_0 < \hat{x} \equiv \sigma_\epsilon^2/(1-\alpha)$ .  
**or** (ii)  $\alpha = 1$ .  
**or** (ii)  $\alpha > 1$  and  $\sigma_r^2 - \sigma_\epsilon^2/(\alpha-1) < 0$ .

**Proof:** This straightforward but lengthy proof is available on request. ■

Figure A.1 is drawn under condition (M) and shows how the productivity,  $Eq_n$ , on a fixed grade  $n$  varies as the research grade  $k$  varies. Condition (M) ensures that the myopic firm will optimally choose to research that grade on which it is about to produce.

**Proof of Theorem 6.B.1:** Consider  $x_1$  as fixed. Then from (28) we may consider  $x'_N$  a function,  $F$  say, of the sequence of effective signals:  $x'_N = F(R_1, \dots, R_N)$ . Fix an  $R$ . Let  $\bar{x}$  be the fixed point of the map  $h_2(h_1^R(\cdot))$ , so that  $\bar{x} = h_2(h_1^R(\bar{x}))$ . Hence when  $\bar{x}_1 = \bar{x}$  and  $R_m = R \forall m$ ,  $F(\cdot) = \bar{x}$ ; when  $\bar{x}_1$  is close to  $\bar{x}$ , so too will  $F(\cdot)$  be close to  $\bar{x}$ . The Taylor series expansion of  $\bar{x}_N$  around  $R_m = R$  for  $m=1, \dots, N$ , is therefore:

$$\bar{x}_N = \sum_{m=1}^N (\partial \bar{x}_N / \partial R_m) R_m + C \quad \text{where}$$

$$(39) \quad C = \bar{x} - \sum_{m=1}^N (\partial \bar{x}_N / \partial R_m) R$$

and where  $\partial \bar{x}_N / \partial R_m$  is evaluated at  $R_m = R$  for  $m=1, \dots, N$ .

Given any function  $f(x)$  of  $x$  let  $f'(x) \equiv \partial f(x) / \partial x$ ; i.e., a prime denotes a derivative with respect to the  $x$  variable. Define

$$A \equiv -\partial h_1^r(\bar{x}) / \partial \tau \quad \text{evaluated at } \tau = R \quad \text{and}$$

$$\delta \equiv \alpha h_1^{R'}(\bar{x}).$$

Then it is easy to check that when evaluated at  $R_m = R$  for  $m=1, \dots, N$  and at  $x_1 = \bar{x}$ ,

$$\partial \bar{x}_N / \partial R_m = -A \delta^{N-m}.$$

It therefore remains only to find expressions for  $A$  and  $\delta$ . It is easy to verify that

$$A \equiv -\partial h_1^r(\bar{x}) / \partial \tau = [h_1^r(\bar{x})]^2 / \sigma^2 = \text{and } \delta \equiv \alpha h_1^{R'}(\bar{x}) = \alpha [h_1^R(\bar{x}) / \bar{x}]^2.$$

The first equality in Theorem 6.B.1 therefore follows from this and (39). The second equality in Theorem 6.B.1 follows from the definition of  $R_m$ .

It therefore remains only to show that  $\delta < 1$ . First, since  $h_1^R(\tilde{x}) < \tilde{x}$ ,  $\delta < 1$  whenever  $\alpha < 1$ . Next, since  $\tilde{x}$  is a fixed point,  $\tilde{x} = \alpha h_1^R(\tilde{x}) + \sigma_\epsilon^2$  so  $h_1^R(\tilde{x}) = (1/\alpha)[\tilde{x} - \sigma_\epsilon^2]$ . Hence  $\delta = \alpha h_1^R(\tilde{x})^2 / \tilde{x}^2 = (1/\alpha)\{[\tilde{x} - \sigma_\epsilon^2] / \tilde{x}\}^2$ . Since  $\tilde{x} - \sigma_\epsilon^2 \geq h_2(0) - \sigma_\epsilon^2 > 0$ , this latter expression is less than one whenever  $\alpha < 1$ . ■

**Proof of Theorem 6.C.2:** If all the future signals are concentrated on one grade, grade  $N+k$ , we may write with abuse of notation  $x_N'' = g(x_N', \rho_{N+k})$ . A linear expansion of this around some  $(\tilde{x}, \tilde{\rho})$  results in:

$$x_N'' = c_1 x_N' - c_2 \rho_{N+k} + \tilde{c}$$

where

$$\begin{aligned} c_1 &= \partial g / \partial x, \quad c_2 = -\partial g / \partial \rho \quad \text{both evaluated at } (x, \rho) = (\tilde{x}, \tilde{\rho}) \quad \text{and} \\ \tilde{c} &= g(\tilde{x}, \tilde{\rho}) - c_1 \tilde{x} + c_2 \tilde{\rho}. \end{aligned}$$

By definition the capital without future signals is  $K_t^* = 1 - \sigma_w^2 - x_N'$ . Hence  $x_N' = 1 - \sigma_w^2 - K_t^*$ . The research capital with *future* signals is therefore

$$\begin{aligned} K_t^{**} &= 1 - \sigma_w^2 - x_N'' = 1 - \sigma_w^2 - \{c_1(1 - \sigma_w^2 - K_t^*) - c_2 \rho_{N+k} + \tilde{c}\} \\ &= c_1 K_t^* + c_2 \rho_{N+k} + c_3 \end{aligned}$$

where  $c_3 = (1 - c_1)(1 - \sigma_w^2) - \tilde{c}$ . This proves Theorem 6.C.2. ■

**Proof of Proposition 7.A.1.:** Since it depends only on  $\sigma_w^2$  (which is the same for the two firms) the map  $h_1$  does not depend on the grid size. But the  $h_2(x)$  map does. Performing  $u$  iterations of  $h_2$  from  $x$ , and calling the resulting map  $h_2^u(x; B)$ , gives, for the parameters faced by firm B,

$$h_2^u(x; B) = \alpha_B^u x + [(1 - \alpha_B^u) / (1 - \alpha_B)] \sigma_{B\epsilon}^2$$

This is how firm B's  $x$  changes when it upgrades by  $u$  steps. For firm A, on the other hand,  $h_2(x; A) = \alpha_A x + \sigma_{A\epsilon}^2$ . Given the relation between the two sets of parameters assumed in the proposition, we see that for each  $x$ ,  $h_2(x; A) = h_2^u(x; B)$ . And since the map  $h_1$  is the same for them both, their  $(x_t)$  sequences will be identical if they have the same set of effective signals. ■

**Proof of Proposition 7.B.2:** (i): In (31),  $h_1$  does not depend on  $u$ . Also,  $h_1(x) \leq x$ , and  $h_2(x) < \hat{x}$  whenever  $x < \hat{x}$ . Hence the argument in each of the  $h_2$  functions on the right-hand-side of (31) is less than  $\hat{x}$ . Then since  $h_2^u(x) = \alpha^u x + [(1-\alpha^u)/(1-\alpha)]\sigma_\epsilon^2$  is increasing in  $u$  whenever  $x < \hat{x}$ ,  $x'_N$  must be increasing in  $u$ .

(ii) The parameters  $a$  and  $b$  are the relative weights on the learning by doing versus the research signals. These are independent of  $\alpha$  and  $\sigma_\epsilon^2$ , and hence independent of  $u$ .

(iii) The formula for  $\delta$  is

$$(40) \quad \delta \equiv \alpha^u h_1^R(\tilde{x}_u)^2 / \tilde{x}_u^2 = \alpha^u [\sigma_w^2 / (\sigma_w^2 + R\tilde{x}_u)]^2 .$$

Now,  $\tilde{x}_u$  is the fixed point of the function  $h_2^u(h_1(x))$ . This is a strictly increasing and strictly concave function of  $x$  which is strictly positive at  $x=0$ . Further,  $h_2^u(h_1(\hat{x})) < \hat{x}$ . Hence  $\tilde{x}_u$  lies in  $(0, \hat{x})$ . Also,  $h_2^u(x)$  is increasing in  $u$  for each  $x \in [0, \hat{x}]$ . Hence  $\tilde{x}_u$  is strictly increasing in  $u$ . (Indeed, one can prove that  $\tilde{x}_u$  increases monotonically to  $\hat{x}$ .) Hence from (40) we conclude that when  $\alpha < 1$ ,  $\delta$  is decreasing in  $u$ . ■