

The Adoption and Adaptation of Externally Originated Ideas

Fidel Perez-Sebastian*

Universidad Pública de Navarra. Departamento de Economía. 31006 Pamplona

E-mail: fidel.perez@unavarra.es

Revised: February, 1998

Abstract

This is a vintage-capital model of technology adoption that takes seriously the fact that new inventions are specific to the environment in which they emerge. The key feature of the model is that the firm can invest resources in R&D to adapt externally originated ideas to the environment in which they are used. We show that because of the possibility of investing in R&D, differences between the inventor's and user's environments can explain why some firms invest in old techniques. Several predictions of the model are consistent with observation.

Key words: R&D; Adoption; Adaptation; Optimal timing

JEL Classification: D92; O33

*I would like to thank B. Ravikumar and Aubhik Khan for their constant encouragement and advice in writing this paper. I am also grateful to Andrew John, Robert King, Alex Wolman, two anonymous referees, and seminar participants at the University of Virginia, Universidad Carlos III, Universidad de Alicante, and Institut d'Anàlisi Econòmica for useful comments.

1 Introduction

Historical evidence shows that technology characteristics critically depend on the particular set of inputs and their relative prices (hereafter, the “environment”) within which the technology emerges (see Solo (1966) and Rosenberg (1976, 1982)). However, once it exists, technology is also used outside the inventor’s environment. Moreover, there is strong empirical evidence that externally originated ideas are in fact the most important source of productivity growth for most firms (Eaton and Kortum (1996), and Coe et al. (1997)). The evidence points out as well that firms that use a new idea under conditions that are different from the ones that generated it usually modify the new technique to better adapt it to their specific environments, and increase the idea’s value.¹ But adaptation is costly. Information needed to properly modify the new ideas cannot be acquired automatically through repetitive experience in production; it requires the specialization of some of the firm’s resources on deliberate R&D effort.² Furthermore, these R&D expenses constitute a significant part of total adoption cost. For example, Ozawa (1966) estimates that one-third of Japan’s R&D expenditures over the period 1957-62 was spent adapting imported technology to Japanese requirements. Teece (1977) uses a sample of 26 international transfers of technology to show that, on average, the cost of transmitting, absorbing, and adapting the new knowledge accounts for 19 percent of total project costs.

This paper is an attempt to understand the role of differences between the inventor’s and the user’s environments in explaining the different technology adoption experiences observed across firms. We build a model of technology adoption that can answer the following questions. Why do firms use old ideas even though apparently

¹A good example is the Japanese experience. Rosenberg (1982, p. 271) writes: “In a variety of ways, the Japanese adapted Western technologies so as to reduce the capital-output ratio. In textiles, for example, they purchased older, secondhand machines ... once installed, they operated the machinery at higher speeds and for longer hours than in England or America, and they lavished greater amounts of labor in servicing the machines and maintaining them. When the Japanese eventually built their own textile machines, they substituted wood for iron wherever possible. They introduced cheaper raw materials into production, as in the case of cotton spinning, and then added more labor to each spinning machine.”

²Aw and Batra (1994) find empirical evidence that Taiwanese manufacturing firms that commit resources to modifying the acquired technology have higher average productivity levels than those that have no such resource commitments. See also Dahlman and Sercovich (1984).

more productive ideas are available?³ Why some firms adopt new techniques more often than other firms? What are the determinants of R&D investment associated with technology adoption? In our framework, ideas carry a parameter that proxies the difference between the firm's and the investor's environments. R&D effort affects this parameter, reducing the risk associated with the idea currently used in production. In each period, the representative firm must decide whether or not to adopt a new production idea, and the amounts of resources to allocate to R&D and production. Newer techniques are more productive, but the firm loses all the risk reduction achieved with the old technique when a new one is adopted. We study dynamics under the constraint analyzed by Chari and Hopenhayn (1991): the firm can adopt any idea, but only up to some frontier that advances exogenously.

Succar (1987) endogenizes the assimilation process of foreign ideas through R&D. But her model does not analyze the firm's adoption decision. R&D effort plays an important role in Helpman and Trajtenberg (1994). In contrast to the last work, in our framework a new idea arrives in every period, and the firm selectively chooses among the available techniques. Uncertainty about the profitability of new techniques is also present in Jensen (1982). Unlike us, Jensen considers that the firm learns about a new technique in a Bayesian fashion using costless, exogenously-given information.

Rothschild (1971) develops a model similar to ours. He studies how the firm's capital investment decision is affected by changes in market conditions when there are adjustment costs. He however does not allow the firm to internalize future changes; thus the firm follows a myopic behavior. Another difference is that we focus on technology adoption, given specific foundations to the kind of costs that we address.⁴

As ours, Parente (1994), Jovanovic and Nyarko (1996), Chari and Hopenhayn (1991) and Farzin et al. (1998) present vintage capital models of technology adoption. Parente, and Jovanovic and Nyarko analyze the adoption behavior when there is learning-by-doing. Farzin et al. study the case when the firm faces a stochastic innovation process with uncertainty about both the speed of the arrival and the degree of improvement of new technologies. Chari and Hopenhayn show that complemen-

³See, for example, Nabseth and Ray (1974).

⁴Adjustment costs are a narrower concept than adaptation costs. For example, the former refer to resources devoted to integrating new machinery into a production plant, whereas the latter also includes the costs of changing some of the features and components of the new machinery.

tarities between old and new assets can make firms use old ideas. In our model there is no learning-by-doing, ideas are environment-specific, and R&D effort is critical to the adoption process. That is, the key features of the problem that we address are absent from those models.⁵

We find that the possibility of adapting the technology can lead firms located in environments that show relatively large differences with respect to the inventor's to continue an old idea for several periods, even if new ideas that are more productive in the inventor's environment are available. This occurs when the average productivity growth from adopting new inventions is relatively small compared to the gains in output from R&D investment. The model also predicts that firms with different average productivity levels can show identical adoption patterns.

We show that the firm decides to adopt ideas more often as the difference between the two environments decreases. Adoption of new ideas also occurs more often, and is less costly, as average R&D productivity rises. Our analysis shows as well that R&D effort decreases with average R&D productivity. These predictions of the theory can help explain the wide range of transfer costs associated with the adoption of foreign ideas (between 2 to 59 percent of total adoption costs) found by Teece (1977), and are consistent with the empirical evidence.

The remainder of the paper is organized as follows. The firm's problem is described in section 2. Section 3 presents the results, and uses them to interpret the adoption patterns displayed last century by U.S. and British steel producers. The final section concludes. The proofs of the Propositions are contained in the Appendix.

2 The Model

2.1 The Production and R&D Technologies

An infinitely lived risk-averse consumer-producer (from now on, "the firm") has preferences given by the CES utility function

$$u(c_t) = c_t^\sigma, \quad 0 < \sigma < 1; \tag{1}$$

⁵The literature on technology adoption is very extensive. We have concentrated on the papers that we believe are most related to our research. For general surveys of models dealing with the optimal timing of adoption of new ideas, see Bridges et al. (1991). Our research is also related to the literature on appropriate technology and growth; see Zeira (1998) and Temple (1998).

where c_t is the amount of consumption at date t . He is endowed with one unit of labor at the beginning of each period. This unit of labor must be allocated between two different activities: production and R&D. The firm produces output using labor, fixed inputs (e.g., land), and one of several ideas indexed by $n = 1, 2, \dots, N$; where N is the “frontier idea” – the last technique that has arrived in the economy. A new idea exogenously arrives at each date. If the firm uses idea n , output in period t (y_t) is given by the following production function:

$$y_t = \gamma^n (\lambda^\theta l_{yt}^\beta), \quad 0 < \beta \leq 1, \quad \theta \in \{\beta, 1\}, \quad \gamma > 1; \quad (2)$$

where l_{yt} is the labor input.

The parameter λ is an environment-specific random variable that depends on uncontrollable events that occur during the production process (e.g., weather variations, and human mistakes), and is a draw from a three point distribution. In particular, if the firm employs idea n at date t , λ is defined on the set $\{m_o - \mu_{nt}, m_o, m_o + \mu_{nt}\}$. We assume that $\text{prob}[\lambda = m_o - \mu_{nt}] = \text{prob}[\lambda = m_o + \mu_{nt}] = p$ for all n and t . Additionally, if idea n has never been used in production before t , the dispersion parameter μ_{nt} equals μ_o ; with $m_o \geq \mu_o > 0$. The parameter θ qualifies how environment differences affect productivity. If θ takes on 1, λ represents an environment-specific component of total factor productivity (TFP). When θ equals β , on the other hand, λ measures the effectiveness of labor or, in other words, the adaptability of the labor force to new ideas.

Differences between the inventor’s and the user’s environments (hereafter, “the environmental difference”) can either affect the mean of λ or its variance or both. For example, modern high-yielding seed varieties developed in the North will be more susceptible to diseases and climate variations in tropical environments; thus increasing output variance. Because the environmental difference equally affects the productivity of all inputs, the example belongs to the $\theta = 1$ case. An adaptation of the high-yielding seed varieties will consist of making them more resistant to the shocks. In this example, uncertainty increases with the environmental difference, and adapting the new idea to the user’s environment reduces this uncertainty.

Other example. Technologies developed in the skill abundant North will be skill-complementary. But the skill level of the labor force in less developed countries is

lower. We are now in the $\theta = \beta$ case: environment differences affect the effectiveness of labor. Adapting the new technology could involve developing a more automated industrial process in which the need for skills would be minimal. In this example, both differences in environment and adaptation can affect either the mean or the variance of λ .

As will be clear later, the firm's adoption decision depends on the ratio of the variance to the mean of λ , and not on their absolute values. Given this, we choose the parameter μ_o to capture the environmental difference. The larger is μ_o , the greater the environmental difference. We assume that environment characteristics remain constant over time.

Our model is one of deliberate and costly information production. R&D effort produces information that allows to reduce the variance of the parameter λ . However, R&D only affects the uncertainty associated with the technique employed by the firm. Specifically, if the firm is currently using idea n , and invests labor in R&D in period t , at date $t + 1$

$$\mu_{n,t+1} = \mu_{nt}(1 - l_{\lambda t}^{\phi}), \quad 0 < \phi < 1; \quad (3)$$

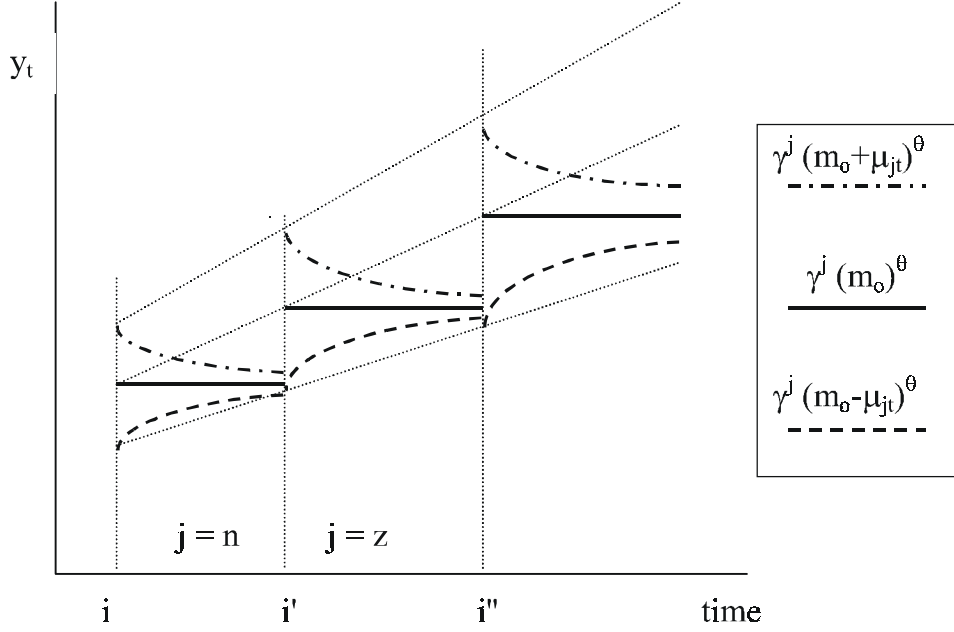
where $l_{\lambda t}$ is the amount of labor devoted to R&D at date t .

Figure 1 illustrates a feasible time path of the output's three point support for $l_{yt} = 1$. Assume that at date i the firm adopts idea n . From i to i' , the firm will use idea n to produce output, and will invest in R&D to reduce the value of the dispersion coefficient μ_{nt} , thus compressing λ 's support around the constant mean m_o . If in period i' a new idea z is adopted, with $z > n$, the support will spread out. Compared to idea n , technique z will be γ^{z-n} times more productive at each value of the dispersion parameter. This pattern will repeat itself in subsequent adoptions.

2.2 The Firm's Problem

At the beginning of period t , given the variance of λ , the firm decides which idea to use, and allocates labor to production and R&D so as to maximize the expected discounted utility over the infinite horizon. If n was the idea used in production at date $t - 1$, the sequence of events in period t is as follows:

Figure 1: time path of the output's support for $l_{yt} = 1$



- If idea n is continued then
 1. Given μ_{nt} , the decision variables $l_{\lambda t}$ and l_{yt} are chosen.
 2. Output is produced, and R&D effort reduces the variance of λ , yielding $\mu_{n,t+1} = \mu_{nt}(1 - l_{\lambda t}^\phi)$.
- If idea $n + i$ is adopted, with $i > 0$:
 1. Given μ_o , the decision variables $l_{\lambda t}$ and l_{yt} are chosen.
 2. Output is produced, and R&D effort reduces the variance of λ , yielding $\mu_{n+i,t+1} = \mu_o(1 - l_{\lambda t}^\phi)$.

The firm's problem can be simplified. Let the technique employed by the firm in the previous period be the "status quo" idea. Since R&D investment only decreases the uncertainty associated with the idea currently employed in production, the expected value of previously abandoned or skipped ideas do not change. Therefore, the firm will never recall them. Given that the level of uncertainty is the same across those ideas that have not been used, the firm will always choose among them the

technique with the highest average productivity, that is, the frontier idea. This is the content of Proposition 1.

Proposition 1 *The firm will decide either to continue the status quo idea or to adopt the frontier idea.*

Next, the firm's problem is stated using a recursive approach. Recall that the firm is a consumer-producer; hence, $u(c_t) = u(y_t)$. From equations (1), (2) and (3), and Proposition 1, if n is the status quo idea, the firm solves the following problem at the beginning of date t :

$$v_n(\tau_{nt}, \mu_{nt}) = \max \{s_n(\tau_{nt}, \mu_{nt}), s_N(0, \mu_o)\}, \quad (\text{P1})$$

where

$$s_n(\tau_{nt}, \mu_{nt}) = \max_{l_{\lambda t} \in [0,1]} \left\{ \gamma^{n\sigma} E_t \left([\lambda(\mu_{nt})]^{\theta\sigma} \right) (1 - l_{\lambda t})^{\beta\sigma} + \delta v_n(\tau_{nt} + 1, \mu_{n,t+1}) \right\}$$

$$s.t. : E_t \left([\lambda(\mu_{nt})]^{\beta\sigma} \right) = p \left[(m_o - \mu_{nt})^{\beta\sigma} + (m_o + \mu_{nt})^{\beta\sigma} \right] + (1 - 2p)m_o^{\beta\sigma}$$

$$\mu_{n,t+1} = \mu_{nt}(1 - l_{\lambda t}^{\phi})$$

$$l_{\lambda t} + l_{yt} = 1;$$

where $\delta \in (0, 1)$ is the discount factor; $v_n(\cdot)$ is the value function associated with idea n ; τ_{nt} is the number of periods since the arrival of invention n or, in other words, the age of idea n at date t ; and $s_n(\cdot)$ is the discounted sum of expected utility when the firm chooses idea n in period t and the optimal policy is followed thereafter.

In order to ensure that the value function is bounded, we impose the following constraint on the discount factor

$$(A1) \quad \delta < 1/\gamma.$$

Given assumption A1, standard dynamic programming arguments ensure that both functional equation (P1) and the maximization problem defined over the discounted sum of expected utility generate the same value function and solutions.⁶

⁶It follows, for example, from Theorems 4.3 and 4.5 in Stokey et al. (1989).

Proposition 2 characterizes the value function $v_n(\cdot)$. It is easy to see that $v_n(\cdot)$ is proportional to $(\gamma^n m_o^\theta)^\sigma$, and that the evolution of the state variables does not depend on the numerical values of the idea indices. The solution to (P1) then will not depend on n or on calendar time; it will depend on the ratio of μ_{nt} to m_o , and on the status quo idea age. In that sense, the firm's problem is stationary.⁷

Proposition 2 (*Stationarity*). (i) *The optimal policy is independent of the numerical value of the status quo idea, and depends only upon the ratio of the dispersion parameter associated with the status quo idea to m_o , and the status quo idea's age.* (ii) $v_{n+1}(\tau, \mu) = \gamma^\sigma v_n(\tau, \mu)$. (iii) $s_{n+1}(\tau, \mu) = \gamma^\sigma s_n(\tau, \mu)$.

Given this result, I assume without loss of generality that $n = 1$ is the status quo idea, and that m_o equals 1. The firm then decides either to continue technique one (from now on, action "CONTINUE") or adopt the frontier idea (from now on, action "ADOPT"). Notice that the frontier idea index equals the age of idea one. Under Proposition 2, problem (P1) can be rewritten as

$$v_1(\tau, \mu) = \max \{s_1(\tau, \mu), \gamma^{\sigma\tau} s_1(0, \mu_o)\}, \quad (\text{P2})$$

where

$$s_1(\tau, \mu) = \max_{l_\lambda \in [0,1]} \left\{ \gamma^\sigma \left(p \left[(1 - \mu)^{\theta\sigma} + (1 + \mu)^{\theta\sigma} \right] + (1 - 2p) \right) (1 - l_\lambda)^{\beta\sigma} + \delta v_1 \left[\tau + 1, \mu \left(1 - l_\lambda^\phi \right) \right] \right\}.$$

Since we have normalized m_o to 1, μ_o and μ_{nt} are now measured in terms of m_o .

Proposition 3 (i) $s_1(\tau, \mu)$ is nondecreasing in τ , strictly decreasing in μ , bounded, and continuous. (ii) $v_1(\tau, \mu)$ is nondecreasing in τ , nonincreasing in μ , bounded, and continuous. (iii) *The set of optimal policies is nonempty.*

3 Optimal Dynamic Paths

In this section, we identify the optimal policies for different parameterizations of the model. Proposition 4 states the technology-switching rule. Proposition 5 describes

⁷Clearly, the solution will also depend on the parameters γ , δ , σ , θ and β .

the optimal policy. Propositions 6 and 7 and the simulation exercise present the main results.

The firm will adopt a new technology if and only if the discounted sum of expected utility is higher with that idea than with the status quo. It turns out that this occurs if and only if the dispersion parameter takes on a value higher than some threshold level. When the dispersion parameter is sufficiently low with the status quo idea, the gains in current expected output from continuing the same technique may offset the expected benefits from adopting the frontier.

Proposition 4 *Fix both the environmental difference μ_o and the age of idea one, τ . The action ADOPT will be chosen over the action CONTINUE if and only if μ is greater than the threshold level $\hat{\mu} > 0$; where $\hat{\mu}$ is given by $s_1(\tau, \hat{\mu}) = \gamma^{\sigma\tau} s_1(0, \mu_o)$.*

The firm will never continue an idea an infinite number of periods because average productivity grows without bound with the frontier idea, whereas expected output from continuing the same technique an infinite number of periods is finite; at some point, adopting the frontier will be strictly preferred. Moreover, the firm's behavior is stationary. From Proposition 2, neither the adoption decision nor the labor allocation depend on the status quo idea's numerical index. Therefore, the optimal policy between consecutive dates at which the firm adopts new ideas will always be the same: the firm will continue an idea a finite number of periods x , then will adopt a new one that will again be continued x periods, and so on. With respect to the R&D allocation, it will remain between zero and one if the firm is continuing the status quo idea next period because the marginal productivity of labor (MPL) explodes as labor goes to zero. Otherwise, R&D effort will be zero because R&D raises productivity only if the firm continues the modified technique in subsequent periods.

Proposition 5 *The optimal policy is to take action CONTINUE for $\hat{\tau}$ consecutive periods, with $\hat{\tau} \geq 0$, then take action ADOPT, then take action CONTINUE for $\hat{\tau}$ periods, and so on. If the action CONTINUE is to be taken next period, the optimal l_λ is greater than zero, but less than one; l_λ equals zero otherwise.*

We are now ready to analyze the determinants of the firm's adoption decision. It is straightforward that adopting a new idea at each date is optimal for a sufficiently low

discount factor, regardless of the environmental difference; notice that if the firm is myopic (i.e., $\delta = 0$), R&D will have no value. Proposition 6 says that another sufficient condition for the firm to adopt a new idea in every period is to have a sufficiently large productivity growth rate γ . The key is that the benefits from continuing an idea are finite, whereas the gains from adopting new ideas grows without limit with the productivity growth rate. From Proposition 5, if the firm always decides to adopt new ideas, R&D investment will never take place.

Proposition 6 *Fix the environmental difference μ_o . There exists $\hat{\gamma} > 1$ such that for all $\gamma > \hat{\gamma}$ the optimal policy is to choose the action ADOPT in every period, and a zero R&D investment.*

The increase in average productivity that the adoption of a new idea brings about is then critical in explaining rapid adoption episodes, and must be incorporated in empirical research as an important explanatory variable of the firm's adoption decision. Its inclusion might help to reconcile some empirical results. For example, Lücke (1993) finds that a country's level of economic development is positively correlated with the date at which the first decision to adopt new ideas is made in this country's textile industry. He however does not find evidence of correlation in the steel industry. If the adoption of the new technologies considered in his research generates a larger increase in average productivity in the steel industry than in the textile industry, this would provide a possible explanation of Lücke's findings.

Proposition 7 illustrates the importance of the environmental difference. While a low-difference firm will always adopt new ideas at each date, a high-difference firm may prefer to continue the status quo idea for several periods before adopting a new one. Using a new technique in every period is not optimal if the average productivity growth rate from adopting new ideas is sufficiently small, and the discounted returns to R&D are sufficiently large. The production and utility functions play an important role in the firm's adoption decision. Returns to R&D are high enough only if uncertainty sufficiently reduces expected utility. In turn, this occurs if either the production function displays strong enough diminishing returns to effective labor or the agent is sufficiently risk averse.

Proposition 7 (i) *There exist $\hat{\mu}_L, \hat{\mu}_U \in (0, m_o)$, with $\hat{\mu}_L \leq \hat{\mu}_U$, such that for all $\mu_o < \hat{\mu}_L$ the optimal policy is to take the action *ADOPT* in every period; and for all $\mu_o > \hat{\mu}_U$ there exist \hat{p} and $\hat{\delta}$ sufficiently big, and $\hat{\gamma}$ and $\hat{\theta}\hat{\sigma}$ sufficiently small, such that for all $p \in (\hat{p}, \frac{1}{2}]$, $\delta \in (\hat{\delta}, 1)$, $\gamma \in (0, \hat{\gamma})$, and $\theta\sigma \in [\frac{1}{2}, \hat{\theta}\hat{\sigma})$ the optimal policy involves taking the action *CONTINUE* a strictly positive number of periods before choosing the action *ADOPT*.*

In our dynamic setting, differences in environment are then necessary, but not sufficient to make firms invest in old ideas. It is the possibility of substantially increasing output by investing in R&D that causes large-difference firms to keep using for several periods ideas that are old, and less productive in the inventor's environment. Without R&D the firm would always decide to adopt new technologies, regardless of the environmental difference. Since the environmental difference was redefined in terms m_o , Propositions 7 also implies that firms with very different average productivity levels (but with the same μ_o to m_o ratio) will show identical adoption patterns.

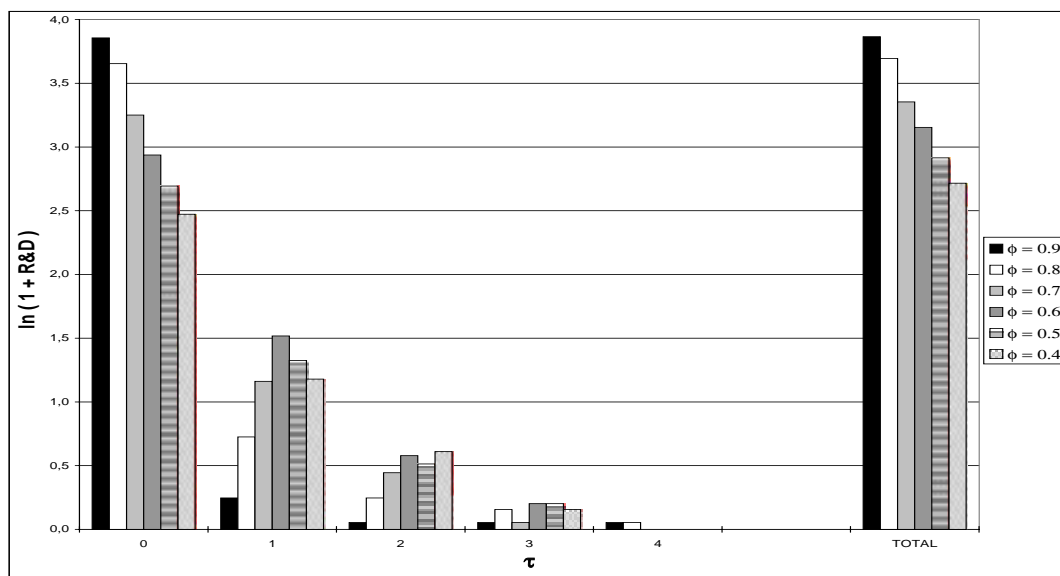
Obtaining closed form solutions to the dynamic programming problem is not possible. So, to better understand the effect on R&D investment of changes in the model parameters' values, we produce simulation results.

3.1 Simulation Exercise

For the numerical analysis, we solved the dynamic programming problem given in (P2) using value function iterations.⁸ We assign a value of 0.95 to the discount factor (δ), and a value of 1.02 to the technological progress parameter (γ), which is the average gross rate of growth of per capita output in the U.S. over the 1950-1985 period. With respect to the parameters θ , ϕ and μ_o , we conduct a sensitivity analysis, and simulate the model for $\theta = \beta$, $\theta, \phi \in \{0.4, 0.5, \dots, 0.9\}$, $\sigma = 1$ and $\mu_o \in \{\frac{1}{20}, \frac{2}{20}, \dots, \frac{19}{20}, 1\}$. It turns out that the (qualitative) effect of changes in one of these parameters is in general insensitive to the values of the other two. For this

⁸The state space has two dimensions. The idea age, τ , takes only discrete values, but the dispersion coefficient μ is a continuous variable. To carry out the simulation exercise then, we assume that $\mu \in \{\frac{1\mu_o}{400}, \frac{2\mu_o}{400}, \dots, \frac{399\mu_o}{400}, \mu_o\}$. Note that in our model μ is always greater than zero.

Figure 2: R&D investment for different values of ϕ ($\beta = 0.6$ and $\mu_o = 0.9$)



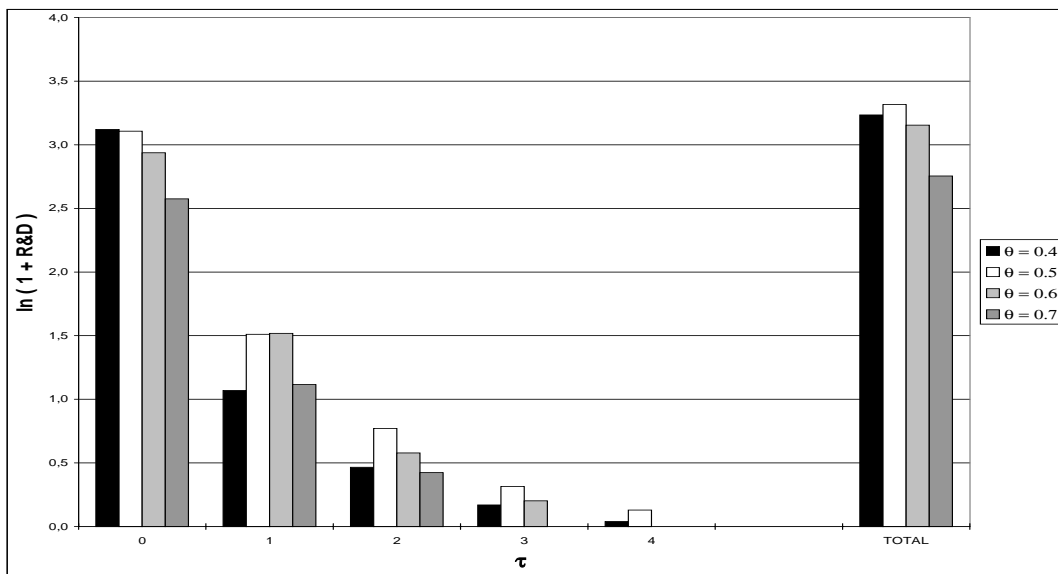
reason, Figures 2 to 4 present the results for selected parameter values. The analysis also allows to understand the effect of changes in σ .

In Figure 2, we show how R&D investment changes with ϕ .⁹ The first striking finding is that over 70 percent of total R&D takes place in the period at which the idea is adopted. It occurs because the rate of return to R&D rapidly declines with μ ; that is, the model dynamics introduce a non-convexity in the cost of adoption. This “spike” investment pattern is not implausible; similar patterns of investment in equipment and machinery have been found in U.S. manufacturing plants.¹⁰ Second, the model predicts that total R&D effort rapidly decreases with average R&D productivity (i.e., as ϕ declines, since labor is bounded between *zero* and *one*). In Figure 2, the total amount of resources devoted to R&D when ϕ equals 0.9 is over three times the R&D investment when ϕ is 0.4. The key now is that the amount of resources needed to reach

⁹To be able to compare the R&D allocations across time and for different parameters’ values, we weight the amount of labor devoted to R&D by its marginal productivity. The resulting amount is the R&D investment measure. In addition to this, to understand the results presented in Figures 2 to 4, notice that the vertical axis gives $\ln(1 + \text{R\&D effort})$. A jump of x units then means that R&D investment experiences a e^x fold increase (this approximation is valid when the natural log takes on a value greater than 2.5).

¹⁰Doms and Dunne (1994) find that US manufacturing plants concentrate about 50 percent of their cumulative 17 year investment in the three years surrounding the year with the largest investment.

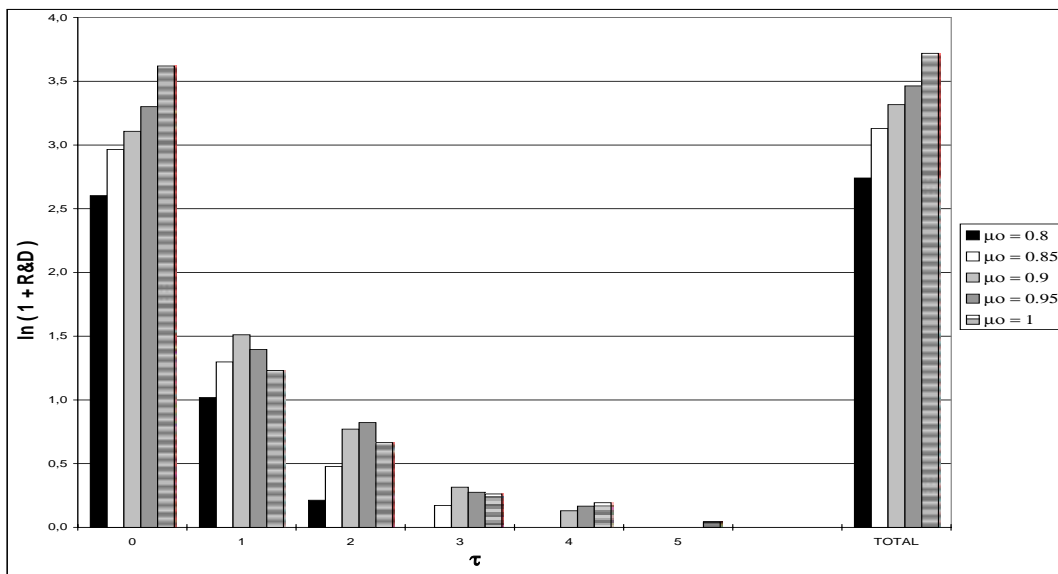
Figure 3: R&D investment for different values of θ ($\phi = 0.6$ and $\mu_o = 0.9$)



any given risk reduction level declines with R&D productivity. Another interesting result is that the firm decides to adopt new ideas more frequently as ϕ declines. For example, the firm continues the status quo idea for 4 periods if $\phi \in \{0.8, 0.9\}$, whereas it is continued only for 3 periods if $\phi \in \{0.4, 0.5, 0.6, 0.7\}$. The intuition is the following. As average R&D productivity rises, expected discounted utility both from continuing the same idea and from adopting a new one go up; but the latter increases faster because it is associated with a (γ times) higher zero-variance productivity. The model can therefore explain why, as R&D productivity goes up, adoption costs decline (Ozawa (1966) and Teece (1977)), and the firm adopts foreign ideas more often (Eaton and Kortum (1995)).

Technology adoption also occurs more often (see Figure 3) as diminishing returns to effective labor get weaker. For example, when θ is 0.4, a new technique is adopted every five periods, whereas the firm never chooses to continue an idea for θ equal to 0.8 and 0.9. When θ goes up, the return from investing in production grows faster than in R&D both because the firm discounts (by a factor of δ) future gains in expected output and because the effect of uncertainty declines as the production function becomes more linear over effective labor. This discourages continuing old

Figure 4: R&D investment for different values of μ_o ($\beta = 0.5$ and $\phi = 0.6$)



techniques. The effect on R&D investment however is not monotonic. The above arguments imply that the total amount of labor allocated to R&D must decline with θ . But the MPL actually rises with θ , which implies that R&D becomes more costly (more output is given up per unit of R&D effort). The effect on the MPL dominates for low values of θ , causing total R&D investment to increase, whereas the effect on R&D productivity prevails for high values, and total R&D investment declines.

We can deduce the effect of varying σ from the results for θ . A rise in σ has the same consequences as an increment in θ . Additionally, the rise in σ increases the TFP gain from adopting new ideas, making action ADOPT more attractive. Hence, if the frequency of adoptions goes up and the amount of resources devoted to R&D declines with θ , the same will occur with σ .

Figure 4 presents the results regarding changes in the environmental differences. We see that the firm decides to adopt new ideas more often and to reduce R&D effort as μ_o declines. For example, total R&D investment almost triples and $\hat{\tau}$ doubles when μ_o goes up from 0.8 to 1. Another variable that can help to explain the wide range of transfer costs reported by Teece (1977) is then the environmental difference.

Besides the investment spike behavior, there is another result that is common to

Figures 2 to 4. The effect of changes in the parameters is not uniformly distributed over the whole interval between adoptions. Larger-difference firms, for instance, invest more in R&D at the time of adoption, but they may choose to invest less in subsequent periods. The same is true for changes in R&D productivity and in the degree of diminishing returns. For example, looking at Figure 2, we see that $\phi = 0.9$ represents the highest R&D investment at the time of adoption, $\phi = 0.6$ does one period later, and $\phi = 0.4$ when τ is two. The reason is that R&D productivity declines with the dispersion parameter. A greater R&D investment today then leads to a lower R&D effort tomorrow. In our model, adaptation-related investment lags cannot then capture firm-specific effects in most cases; cumulative investment is a better candidate to capture them.

3.2 Adoption Patterns in the Iron Industry

The eighteenth and nineteenth centuries saw one of the iron industry's most remarkable achievements: the switch from using charcoal to employing mineral coke to smelt the iron ore. This switch however did not occur at the same time in all countries. Rosenberg (1976) gives a good example comparing firms' behaviors in the U.S. and Britain. During this period of time, a sequence of new ideas were proposed by British inventors to find a cheap substitute for wood fuel (charcoal). Darby developed his coke-smelting process in 1709. In Britain, the shift from charcoal to coke was virtually completed by 1800. But as late as 1840, 100% of pig iron in the U.S. was being smelted using charcoal. The lack of success of the new idea among U.S. firms was related to the poor quality of the U.S. bituminous coke and its specific location, far from the country's population centers. A second idea was to reduce heat losses by using a hot blast.¹¹ It was applied in England in 1828. But again for the same reasons, few U.S. firms adopted the new idea. A Welshman, David Thomas, altered this situation in the 1830's. He developed the anthracite blast furnace. This time, the new invention was widely adopted in the U.S. because the anthracite coal was highly abundant and readily accessible in North America.

Proposition 7 implies that our model can exactly reproduce the above adoption

¹¹The hot blast technique permitted substantial fuel economies by simply preheating the blast before entering the furnace, and later by employing waste gases from the furnace itself.

patterns. British firms adopted all ideas created by British inventors because the environmental difference was small, whereas U.S. firms selectively chose among them because the environmental difference was big. Moreover, we can provide, in terms of our dynamic setup, additional reasons for the adoption delay. U.S. firms could have skipped Darby's coke-smelting process and the hot blast technique because the charcoal blast furnace had already been modified and adapted to the inputs available in the U.S.. Only when the productivity level associated with the frontier idea became sufficiently high, U.S. firms would have decided to adopt the new invention, jumping from using charcoal to installing the anthracite blast furnace. Whether adaptation played an important role is clearly an empirical issue, outside the scope of this paper.

4 Summary and Concluding Remarks

The theory of technology adoption advanced here can explain why some firms invest in old techniques even though apparently more productive ideas are available. The key is the possibility of investing in R&D to adapt externally originated ideas to the environment in which they are used. While adopting a new technique raises average productivity, it also causes the risk reduction achieved with the old technology to be forfeited. In this sense, R&D investment is an *irreversible* decision. We have also shown that difference between the two environments and average R&D productivity critically affect both the optimal timing of adoption and the optimal allocation to R&D.

In our framework, even though adoption costs are convex at each point in time, investment lags do not necessarily give any information about the firm's type. Larger-difference firms invest more in R&D at the time of adoption, but they may invest less in subsequent periods. This result is interesting because it adds new concerns to the ones advanced, for example, by Rothschild (1971) about the use of distributed lags of past investment to capture firm-specific effects in empirical research. Both cumulative investment and total investment per idea arise as better candidates to capture them. The result must, however, be interpreted with caution; it might be the consequence of our simplifying assumptions. In order to believe this finding, it would be important to see whether it holds when we allow the firm to use a technique and meanwhile to

separately search and adapt new ideas. This is left for future research.

Clearly, the theory has also implications for the diffusion of new ideas. In our model, it is easy to show that the existence of R&D spillovers could generate a technology diffusion process across firms located in different environments. This is a typical result in technology adoption models that contain information externalities.

Firms in the model always jump to the technology frontier. The model cannot, therefore, account for the “intermediate” adoption patterns observed across firms. We obtain this result because the benefit from adopting the newest invention is the same regardless of which idea the firm is currently using. It might then be important to endogenize the firm’s ability to move up on the technology ladder to account for the intermediate patterns, and is left for future research.

Appendix. The Proofs¹²

Proof of Proposition 1: Obvious. ■

Proof of Proposition 2: Pick $\mu_{nt} = \mu$, and $\tau_{nt} = \tau$. Note that $N = n + \tau$. Let $(Tv)(\tau, \mu) = \max\{s_n(\tau, \mu), s_{n+\tau}(0, \mu_o)\}$. Let $\bar{v}(\cdot)$, $g(\cdot)$, $\hat{v}(\cdot)$, $\bar{s}(\cdot)$, and $l(\cdot)$ be functions that do not depend on n nor on t nor on m_o . Now, using standard dynamic programming arguments, conjecture that $v_n[\tau + 1, \mu(1 - l_\lambda^\phi)] = (\gamma^n m_o^\theta)^\sigma \bar{v}[\tau + 1, fL(1 - l_\lambda^\phi)]$ or, in other words, that the value function is proportional to $(\gamma^n m_o^\theta)^\sigma$, and that we can substitute the state variable $f * L$ for μ ; where the motion equation of L is $L' = L(1 - l_\lambda^\phi)$; and $f = \frac{\mu_o}{m_o}$. Substituting this conjecture into (P1), we can write $s_n(\tau, \mu) = (\gamma^n m_o^\theta)^\sigma \max_{l_\lambda \in [0, 1]} \{g(\tau, fL; l_\lambda)\}$. The R&D investment optimal policy is given by $l^* = l(\tau, fL)$; and hence $s_n(\tau, \mu) = (\gamma^n m_o^\theta)^\sigma \bar{s}(\tau, fL)$. It is then straightforward that the firm's decision depends on fL and τ , and that $(Tv)(\tau, \mu) = (\gamma^n m_o^\theta)^\sigma \hat{v}(\tau, fL)$.

Since μ and τ were arbitrary, if we show that the conjecture is true, we are done. Assumption A1 implies that $\delta\gamma < 1$. It is then immediate that $(Tv)(\cdot)$ defines a mapping that satisfies Blackwell's sufficient conditions for a contraction. Because the space of functions that depend on a given set of arguments with the sup-norm is obviously a Banach space, it follows from the Contraction Mapping Theorem that the conjecture is true. ■

Proof of Proposition 3: We first show that τ is bounded above; that is, the firm never continues idea one an infinite number of periods. By contradiction, assume that the firm chooses to continue idea one forever. This implies

$$\gamma^{\sigma\tau} s_1(0, \mu_o) \leq \sum_{t=1}^{\infty} \gamma^\sigma \delta^{t-1} E\left([\lambda_o(0)]^{\theta\sigma}\right) = \left(\frac{\gamma^\sigma}{1 - \delta\gamma^\sigma}\right), \quad \text{for all } \tau.$$

Notice that the RHS is finite. However, $\gamma^{\sigma\tau} s_1(0, \mu_o)$ goes to infinity as τ goes to infinity. Therefore, we can find a τ sufficiently large such that $\gamma^{\sigma\tau} s_1(0, \mu_o) > \left(\frac{\gamma^\sigma}{1 - \delta\gamma^\sigma}\right)$; and we got a contradiction.

Let $(Tv)(\tau, \mu) = \max\{s_1(\tau, \mu), \gamma^{\sigma\tau} s_1(0, \mu_o)\}$. Conjecture that $v_1(\tau, \mu)$ is nondecreasing in its first argument, nonincreasing in its second argument, bounded, and continuous. Note that l_λ is defined in a nonempty, constant, and compact interval;

¹²Some of the proofs make use of the Theorem of the Maximum, Blackwell's sufficient conditions for a contraction, and the Contraction Mapping Theorem. These theorems can be found, for instance, in Stokey et al. (1989).

and that $E_t[u(y_t)]$ (first additive term in $s_1(\cdot)$) is strictly decreasing in μ , bounded since τ is bounded and $l_\lambda \in [0, 1]$, and continuous. Therefore, it follows that: (1) from the Theorem of the Maximum, $s_1(\cdot)$ is continuous; (2) the term being maximized in $s_1(\cdot, \mu)$ is uniformly decreasing for all l_λ , and hence $s_1(\cdot, \mu)$ is strictly decreasing; and (3) $s_1(\cdot)$ is bounded since $\mu \in [0, 1]$. We also need to show that $s_1(\tau, \cdot)$ is nondecreasing. This follows directly from the fact that the average productivity associated with the frontier idea increases with the age of idea one. Given these properties of $s_1(\cdot)$, it is immediate that $(Tv)(\cdot)$ preserves the properties of $v_1(\cdot)$ given by (ii).

Our next task is to show that the conjecture about $v_1(\cdot)$ is true. It is a well known result that the space of bounded, continuous, nonincreasing/nondecreasing functions with the sup-norm is a Banach space. Because $(Tv)(\cdot)$ defines a contraction (see Proposition 2), it follows from the Contraction Mapping Theorem that the conjecture is true, and consequently $v_1(\cdot)$ has the desired properties.

Finally, given the above, the Theorem of the Maximum applied to (P1) guarantees that the optimal policy set is nonempty. ■

Proof of Proposition 4: Given the definition of the value function (P2), the proposition follows directly from Proposition 3(i). ■

Proof of Proposition 5: From the proof of Proposition 3, we know that $\hat{\tau}$ is finite. Recall also that when the firm adopts a new idea $\mu_{nt} = \mu_o$. The stationary adoption pattern then follows directly from the last fact and Proposition 2. From equation (2), $\frac{\partial y_t}{\partial l_\lambda} \rightarrow \infty$ as $l_\lambda \rightarrow 1$, and then $l_\lambda < 1$ at the optimum. To see that $l_\lambda > 0$ if the action CONTINUE is to be taken next period, note that $s_1[\tau, \mu(1 - l_\lambda^\phi)] > \gamma^{\sigma\tau} s_1(0, \mu_o)$ must hold under the last choice. Clearly, $\left(\frac{\partial s_1[\tau, \mu(1 - l_\lambda^\phi)]}{\partial l_\lambda}\right) \rightarrow \infty$ as $l_\lambda \rightarrow 0$. Therefore, since $\mu > 0$ because $l_\lambda < 1$ at any date, l_λ will be greater than 0. R&D investment will be zero when the firm is to adopt a new idea next period, R&D effort then has a zero return. ■

Proof of Proposition 6: Pick $\hat{\gamma}$ such that $\max_{l_\lambda \in [0, 1]} \{s_1[\tau, \mu(1 - l_\lambda^\phi)]\} = \hat{\gamma}^{\sigma\tau} s_1(0, \mu_o)$. From Proposition 3, we know that $s_1(\cdot)$ is bounded or, in other words, that the LHS of the equality is finite. Such a $\hat{\gamma}$ then exists because the RHS goes to infinity with $\hat{\gamma}$ for all $\tau \geq 1$. From Proposition 5, it follows that the optimal policy is to take the action ADOPT at each date and $l_\lambda = 0$ for all $\gamma > \hat{\gamma}$. In order to see that $\hat{\gamma} > 1$, note that if $\gamma = 1$, $s_1[\tau, \mu_o(1 - l_\lambda^\phi)] > \gamma^{\sigma\tau} s_1(0, \mu_o)$ for all $l_\lambda > 0$ and τ (see Proposition 3). Under the last inequality, $\left(\frac{\partial s_1[\tau, \mu_o(1 - l_\lambda^\phi)]}{\partial l_\lambda}\right)$ goes to infinity as l_λ goes to zero. The

firm will then choose an l_λ greater than zero, and will continue the status quo idea. Applying continuity, all the above implies that $\hat{\gamma} > 1$. ■

Proof of Proposition 7: If the environmental difference goes to zero, $s_1(\tau, \mu_o) \longrightarrow s_1(\tau, 0)$. At each date, $\gamma^{\sigma\tau} s_1(0, 0)$ defines a value function (P2) upper bound. Hence, $\gamma^{\sigma\tau} s_1(0, 0) > s_1(\tau, 0)$ because the RHS of the inequality implies continuing a less productive idea at least one more period. Since $s_1(\cdot, \mu)$ is continuous, there exists $\hat{\mu}_L \neq 0$ sufficiently small such that, for any $\mu_o \in (0, \hat{\mu}_L)$, $\gamma^{\sigma\tau} s_1(0, \mu_o) > s_1(\tau, \mu_o)$, that is, such that the action ADOPT always dominates.

Next, we prove that the policy “adopt a new idea at each date” is not optimal under the stated conditions. Without loss of generality assume that idea one is used for the first time at date 1. If the firm continues the status quo idea for $\hat{\tau} > 0$ periods

$$s_1(0, \mu_o \mid \hat{\tau} > 0) > \gamma^\sigma (\delta + \delta^2 + \dots + \delta^{\hat{\tau}}) + (\delta\gamma^\sigma)^{\hat{\tau}+1} s_1(0, \mu_o); \quad (4)$$

where the RHS is obtained by letting $l_{\lambda 1} = 1$, and $l_{\lambda t} = 0$ for $t = 2, \dots, \hat{\tau}$. If a new idea is, on the other hand, adopted each period

$$\begin{aligned} s_1(0, \mu_o \mid \hat{\tau} = 0) &= \gamma^\sigma \left(p \left[(1 - \mu_o)^{\theta\sigma} + (1 + \mu_o)^{\theta\sigma} \right] + (1 - 2p) \right) * \\ &* \left[1 + \delta\gamma^\sigma + (\delta\gamma^\sigma)^2 + \dots + (\delta\gamma^\sigma)^{\hat{\tau}} \right] + (\delta\gamma^\sigma)^{\hat{\tau}+1} s_1(0, \mu_o). \end{aligned} \quad (5)$$

Subtracting these expressions we have

$$s_1(0, \mu_o \mid \hat{\tau} > 0) - s_1(0, \mu_o \mid \hat{\tau} = 0) > \gamma^\sigma \left[\delta (1 - \gamma^\sigma \psi) + \dots + \delta^{\hat{\tau}} (1 - \gamma^{\sigma\hat{\tau}} \psi) - \psi \right]; \quad (6)$$

where $\psi = p \left[(1 - \mu_o)^{\theta\sigma} + (1 + \mu_o)^{\theta\sigma} \right] + (1 - 2p)$.

In inequality (6), the RHS decreases with p , μ_o , and rises with γ and $\theta\sigma$ (if $\theta\sigma \geq \frac{1}{2}$). Let $\gamma = 1$, $\theta\sigma = \frac{1}{2}$, $p = \frac{1}{2}$, and $\mu_o = 1$. We can now write expression (6) as follows:

$$s_1(0, \mu_o \mid \hat{\tau} > 0) - s_1(0, \mu_o \mid \hat{\tau} = 0) > (\delta + \delta^2 + \dots + \delta^{\hat{\tau}}) (1 - 2^{-\frac{1}{2}}) - 2^{-\frac{1}{2}}. \quad (7)$$

It is immediate that the RHS of expression (7) is positive if $(\delta + \delta^2 + \dots + \delta^{\hat{\tau}}) > 2.42$. Because $\hat{\tau}$ can be as large as we wish, this will be true for a sufficiently big δ . Applying continuity, we are done. ■

References

- [1] Aw, Bee-Yan, and Geeta Batra (1994). "Technological Capability and Firm Efficiency." Pennsylvania State University Working Paper No.11-94-25.
- [2] Bridges, E., Coughlan, A.T., and Kalish, S., (1991). "New Technology Adoption in an Innovative Market-Place: Micro- and Macro-Level Decision Making Models." *International Journal of Forecasting* 7: 257-270.
- [3] Chari, V.V., and Hugo Hopenhayn (1991). "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy*, vol. 99 (6): 1142-65.
- [4] Coe, David T., Elhanan Helpman (1995) and Alexander W. Hoffmaister, (1997). "North-South R&D Spillovers." *Economic Journal* 197: 134-149.
- [5] Dahlman, Carl J., and Francisco C. Sercovich (1984). "Exports of Technology from Semi-Industrial Economies and Local Technological Development." *Journal of Development Economics* 16: 63-99.
- [6] Doms, M. and T. Dunne (1994). "*Capital Adjustment Patterns in Manufacturing Plants.*" mimeo, Center for Economic Studies, Census Bureau.
- [7] Eaton, Jonathan, and Samuel Kortum (1996). "Trade in Ideas: Patenting and Productivity in the OECD." *Journal of International Economics* 40: 251-278.
- [8] Farzin, Y.H., Huisman, K.J.M., and P.M. Kort (1998). "Optimal Timing of Technology Adoption." *Journal of Economic Dynamics and Control* 22: 779-799.
- [9] Helpman, Elhanan, and Manuel Trajtenberg (1994). "A Time to Sow and a Time to Reap: Growth based on General Purpose Technologies." NBER Working Paper No. 4854.
- [10] Jensen, R., (1982). "Adoption and Diffusion of an Innovation of Uncertain Profitability." *Journal of Economic Theory* 27: 182-193.
- [11] Jovanovic, Boyan, and Yaw Nyarko (1996). "Learning by Doing and the Choice of Technology." *Econometrica* 64 (6): 1299-310.

- [12] Lücke, Matthias, (1993). "The Diffusion of process Innovations in Industrialized and Developing Countries: A Case Study of the World Textile and Steel Industries." *World Development*, 21 (7): 1225-38.
- [13] Nabseth and Ray (1979). *The Diffusion of New Industrial Processes; an International Study*. Cambridge University Press.
- [14] Ozawa, T., (1966). *Imitation, Innovation, and Trade: A study of foreign licensing operations in Japan*. Ph.D. Dissertation Columbia University (New York).
- [15] Parente, Stephen L., (1994). "Technology Adoption, Learning by Doing, and Economic Growth." *Journal of Economic Theory*, 63: 346-69.
- [16] Rosenberg, Nathan, (1976). *Perspectives on Technology*. Cambridge University Press (Cambridge).
- [17] ——— (1982). *Inside the Black Box. Technology and Economics*, Cambridge University Press.
- [18] Rothschild, Michael, (1971). "On the Cost of Adjustment." *Quarterly Journal of Economics* 85: 605-22.
- [19] Solo, Robert A., (1966). "The Capacity to Assimilate an Advanced Technology." *American Economic Review Papers and Proc.* 56: 91-7.
- [20] Stokey, Nancy L., Robert E. Lucas, with Edward C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.
- [21] Succar, Patricia, (1987). "International Technology Transfer: A Model of Endogenous Technological Assimilation." *Journal of Development Economics*: 26: 375-95.
- [22] Teece, D.J., (1977). "Technology Transfer by Multinational Firms: The Resource cost of International Technology Transfer." *Economic Journal*, 87: 242-61.
- [23] Temple, Jonathan, (1998). "Appropriate technology: five growth puzzles solved?" mimeo, University of Oxford (Oxford, U.K.).
- [24] Zeira, Joseph, (1998). "Workers, Machines, and Economic Growth." *Quarterly Journal of Economics* 113(4): 1091-117.