Learning-by-Doing and the Choice of Technology: The Role of Patience

Larry Karp and In Ho Lee[1]

University of California, Berkeley and University of Southampton

July 6, 1998

[1] In Ho Lee acknowledges the ¯nancial support from ESRC Research Fellowship. Remaining errors are ours.

Abstract

Jovanovic and Nyarko (1996) showed that when agents learn-by-doing and are myopic, less advanced agents may adopt new technologies while more advanced ¯rms stick with the old technology since the new technology takes time to learn. In this case, the less advanced agents might eventually overtake (or \leapfrog") the advanced agents. We show that this kind of overtaking can also occur if agents are forward looking and have high discount rates. However, if agents are su±ciently patient, overtaking cannot occur. A lower discount rate increases the set of states at which agents adopt new technologies, so more patient agents tend to upgrade their technology more frequently.

keywords: learning-by-doing, overtaking, leapfrogging, technology adoption, economic growth

JEL Classi¯cation numbers D92, O14,O33

1 Introduction

Modern development economics emphasizes the role of technology in determining relative growth paths. For example, Lucas (1993) identi¯es technology adoption as the most important explanation of the fast economic growth of several Asian countries. Recent textbooks such as Aghion and Howitt (1998) and Barro and Sala-i-Martin (1995) re°ect the importance attributed to technology in explaining growth. Technological improvements can lead to divergence in growth paths when ¯rms in the \advanced" country have a greater incentive to adopt new technology. In other circumstances, ¯rms in less advanced countries may be more likely adopt the new technology, even when it is less pro¯table for them than for the advanced ¯rms. The adoption decision depends on a comparison of pro¯ts under the new technology and under the next best alternative, i.e. on the opportunity cost of adoption. The opportunity cost of adoption may be higher for the more advanced ¯rms, because of their pro¯ciency in using the old technology. In this case, innovations in technology can contribute to the convergence of growth paths, or even to \overtaking" (or \leapfrogging") by the less advanced country.

There have been a number of historical examples where technology adoption has contributed to overtaking, both at the industry and country level. Industries in regions destroyed by war (such as in post-war Europe and Japan) sometimes rebuild using the latest technology, eventually overtaking established industries elsewhere. Start-up industries may begin with the latest technology which incumbents are slow to adopt. Brezis et al. (1993) cite cases where new technologies have contributed to overtaking by entire countries rather than individual sectors: the United Kingdom overtaking the Netherlands, and the United States subsequently overtaking the United Kingdom.

The incentives to adopt a new technology depend on the ¯rm's ability to use the previous generation of technology. This ability may depend on the experience the ¯rm has had with the technology, i..e., on the amount of learning-by-doing that has occurred. Chari and Hopenhayn (1991), Parente (1994) and Stokey (1988) study learning-by-doing as a force for sustained growth. Brezis et al.

(1993), Krussell and Rios-Rull (1996), and Jovanovic and Nyarko (1996) show that learning-by-doing can give rise to the type of overtaking noted in the growth literature. An agent accustomed to an existing technology may be unwilling to adopt a newer technology which requires learning and leads to lower pro¯ts in the short run. An agent who is less familiar with the existing technology has a lower opportunity cost of adopting the new technology. The second agent may adopt the new technology and eventually overtake the ¯rst, who was initially more advanced.[1]

Brezis et al. (1993) study a general equilibrium model in which learning is a non-excludable public good within a country. In this situation, ¯rms have no incentive to consider future payo®s when making their adoption decision. (See their footnote 6.) Jovanovic and Nyarko (1996) study a partial equilibrium model in which learning is internal to the ¯rm, but they assume that ¯rms solve a succession of static problems. Thus, in both of these models, agents are myopic: the adoption decision depends on a comparison of current pro¯ts under the old and new technology.

Forward-looking ¯rms who internalize learning-by-doing would consider the future stream of payo®s in deciding whether to adopt the new technology. In order to determine the sensitivity of the overtaking result to the assumption of myopia, we replace the myopic decision-maker in Jovanovic and Nyarko's (1996; hereafter, JN) model with a forward-looking agent. The possibility of overtaking is robust, in the sense that it can occur even when agents are forward looking. However, overtaking is less likely to occur when agents are forward looking, and it never occurs if agents are su±ciently patient.

We also examine the e®ect of the discount rate on the frequency of adoption. Forward looking agents adopt new technologies more frequently than myopic

[1] There is an industrial organization literature on leapfrogging which is closely related to the economic growth literature we cite in the text. The IO literature emphasizes ¯rms' strategic incentives to change a decision, such as improving technology. Budd et al. (1993) review recent contributions to leapfrogging models in IO, and Brezis et al. (1994) discuss the relation between the two literatures. Motta et al. (1997) study a model in which trade changes a ¯rm's strategic decision (quality, in their case), and overtaking can occur. Their model thus incorporates elements of both the IO and economic growth literatures.

agents. At least for small discount factors, the frequency of adoption is monotonic in the discount factor. Parente (1994) uses simulations to show that adoption occurs more rapidly when firms use higher discount factors, and when capital markets improve. Our results complement these simulations, although the two models are quite different.

2 Model

We adapt JN's model of learning-by-doing by including forward looking agents. The payoff associated with a particular technology depends on an unknown parameter. As the agent learns about this parameter over discrete time, the payoff from the technology increases. An agent working with a technology of grade n chooses x in period t and receives the payoff:

$$q = \gamma^n \left[1 - (y_t - x)^2\right]; \quad \gamma > 1.$$

After observing the payoff, the agent can infer the value of y_t since she knows $\gamma; n; x$. This inferred value is $y_t = \theta_n + w_t$, the sum of two random variables; θ_n is a random variable that depends on the technology grade n_t, and w_t is an i.i.d. normal random variable with zero mean and variance σ_w^2. The agent knows the distribution of w_t. The agent does not know the value of θ_n but has prior beliefs about it. Before learning y_t the agent maximizes the expected payoff by setting x equal to the expected value of y_t, conditional on information available in period t:

$$x = E_t[y_t] = E_t[\theta_n]$$

where the second equality follows from the fact that w_t is white noise.[2] This choice yields the expected payoff:

$$E_t[q] = \gamma^n[1 - \operatorname{var}_t(\theta_n) - \sigma_w^2] \quad (1)$$

where $\operatorname{var}_t(\theta_n)$ is the variance of θ_n conditional on information available in period t:

[2] This solution remains valid in the context of dynamic optimization because the information generation is independent of the action choice.

The agent can switch to a higher grade technology in any period. We assume that technology grades are integer-valued and the agent can switch only to the next grade in one period. There is no cost of switching except that the agent has to learn about the new technology. Skills acquired in working with the old technology are only partially transferable. Different grades of technology are linked to each other according to the following relationship:

$$\theta_{n+1} = \sqrt{\alpha}\,\theta_n + \varepsilon_{n+1} \tag{2}$$

where $\varepsilon_{n+1} \sim N(0; \sigma_\varepsilon^2)$ and θ_n and ε_{n+1} are independent.

The agent updates her prior on θ_n based on the signal y_t: Denote the precision of the unknown technological parameter θ_n in period t by ρ_t and the precision of w_t by γ : $\rho_t = \frac{1}{var_t(\theta_n)}$; $\gamma = \frac{1}{\sigma_w^2}$. Obviously ρ_t and γ can take only positive real values, and we assume that $\gamma > 1$. This restriction implies that agents earn positive profits for sufficiently large ρ. In period 1 the agent begins with a Normal prior on the current technology (the value of θ) with precision ρ_1.

We now describe how ρ changes over time. First suppose that there is no technology switch in period t: Since w_t is a Normal random variable, given the Normal prior on the random variable θ_n, its precision is updated in period t according to the following formula (DeGroot, 1970):

$$\rho_{t+1} = \rho_t + \gamma: \tag{3}$$

If the agent switches to a new technology, the variance is updated through two steps. The first step is due to the technology switch and second to the observation of the outcome from the new technology. The first step transforms the variance (prior to the switch) $var_t(\theta_n)$ to $\alpha \cdot var_t(\theta_n) + \sigma_\varepsilon^2$ (the variance after the switch) due to the transformation of θ_n as in equation (2). The agent then chooses x, observes q_t, infers y_t, and updates beliefs about the value of θ_{n+1}. The second step transforms the post-switch variance using equation (3). Combining the two, the precision in the period after a switch occurs is

$$\rho_{t+1} = \frac{1}{\alpha/\rho_t + \sigma_\varepsilon^2} + \gamma = \frac{\rho_t}{\alpha + \rho_t \sigma_\varepsilon^2} + \gamma: \tag{4}$$

It is convenient to define a function which represents the first step of the updating procedure:

$$h(\pi) = \frac{\pi}{\alpha + \sigma_2^2 \pi}.$$

Hereafter we restrict attention to state space where $\pi \geq h(\pi)$. This restriction is innocuous, since for any initial condition it must be satisfied in finite time, regardless of the agent's upgrade decisions. If the restriction is satisfied at any period, it holds in all subsequent periods. Moreover, given the interpretation of the function $h(\pi)$, the model is sensible only when the restriction is satisfied. (If $\pi < h(\pi)$, upgrading increases precision, which means that the agent knows more about the new technology than about the old technology.)

We now introduce forward looking agents. The agent maximizes the present value of the infinite stream of payoffs with a discount factor, $\beta > 0$. In period 1 the agent starts with an arbitrary grade of technology, which we normalize to be grade $n = 0$. Define $k_t = 0$ if the agent decides to stick with the current technology in period t and $k_t = 1$ if the agent chooses to upgrade. The strategy profile is $(k_1, k_2, \ldots)$. Define $T_n = \min_t \{k_1 + k_2 + \ldots + k_t \geq n\}$, the period in which the agent switches to the nth grade of technology, with the convention that $T_0 = 0$. Given a strategy $(k_1, k_2, \ldots)$, we can use the single-period expected payoff in equation (1) to compute the discounted expected payoff.

Suppose that the agent has the precision π at the beginning of the first period. The sequence problem which maximizes the discounted expected payoff, given precision π and technology grade $n = 0$ is:

$$\begin{aligned} W^*(\pi; 0) &= \max_{(k_1, k_2, \ldots)} W(k_1, k_2, \ldots : \pi; 0) \\ &= \sum_{n=0}^{\infty} \gamma^n \beta^{T_n} \left\{ \sum_{t=T_n}^{T_{n+1}-1} \beta^{t-T_n} \left[1 - \frac{1}{\pi_t}\right] \right\} \end{aligned} \qquad \text{(SP)}$$

where π_t is updated according to either equation (3) or equation (4). The first argument of W^* is the precision, π, and the second is the technology grade, n (here $n = 0$).

We use the sequence problem to formulate the dynamic programming equation (DPE). The payoff from the agent's choice depends on the grade of the technology and the precision. Hence the DPE has two state variables, n_t and η_t:

$$V(\eta_t; n_t) = \max_{k_t \in \{0,1\}} \{F(k_t; \eta_t; n_t) + \beta V(\eta_{t+1}; n_t + k_t)\} \tag{5}$$

where

$$F(k_t; \eta; n) = \begin{cases} \gamma^n[\mu - \frac{1}{\eta}] & \text{if } k_t = 0 \\ \gamma^{n+1}[\mu - \frac{1}{h(\eta)}] & \text{if } k_t = 1; \end{cases}$$

$$\eta_{t+1} = \begin{cases} \eta_t + \nu & \text{if } k_t = 0 \\ h(\eta_t) + \nu & \text{if } k_t = 1; \end{cases}$$

and

$$n_{t+1} = n_t + k_t.$$

An optimal policy, $k^*(\eta; n)$, solves the DPE (5).

3 Preliminaries

We first prove the existence of the solution to the DPE (5) and then show that the optimal adoption decision depends only on the precision, η. The following assumption guarantees the equivalence of the solution to the DPE and the solution to the original sequence problem.

Assumption 1 $\beta\gamma < 1$.

We need the equivalence between the two problems in order to justify using the DPE in later analysis.

Proposition 1 1. There exists a solution to the DPE (5).

2. Under Assumption 1, the solution to the DPE satisfying

$$\lim_{t \to \infty} \beta^t V(\eta_t; n_t) = 0$$

is the unique solution to the sequence problem (SP).

Proof. The proof for part 1 of the proposition is standard; define the operator

$$T\tilde{V} = \max_{k \in \{0,1\}} \{F(k : \eta_t; n_t) + \beta \tilde{V}(\eta_{t+1}; n_t + k)\}.$$

¡From equation (5) and $T\tilde{V}$ is easily seen to be a contraction mapping with modulus β.

The second part follows from the result that if the solution to the sequence problem (SP) is bounded, the solution to the DPE satisfying

$$\lim_{t \to \infty} \beta^t \tilde{V}(\eta_t; n_t) = 0$$

is the unique solution to the sequence problem (SP). (Theorem 4.3 on p.72 of Stokey and Lucas (1989)) Hence it suffices to prove that Assumption 1 implies that the solution to the sequence problem (SP) is bounded:

If $\beta\gamma < 1$, then

$$W^*(\eta; 0) = \max_{(k_1; k_2; \ldots)} W(k_1; k_2; \ldots; \eta_{1;1})$$

$$= \sum_{n=1}^{\infty} \gamma^{n-1} \beta^{T_{n-1}} \left\{ \sum_{t=T_{n-1}}^{T_n - 1} \beta^{t - T_{n-1}} \left[1 - \frac{1}{\eta_t}\right] \right\}$$

$$\leq \sum_{t=1}^{\infty} \beta^t \gamma^{t-1} = \frac{1}{1 - \beta\gamma} < \infty.$$

Therefore the solution to the sequence problem (SP) is bounded. ■

Next we show that the optimal upgrade rule depends on the value of η, but not on the grade of technology n or on time, t.

Proposition 2 The optimal upgrade rule depends only on η. That is, the solution to the DPE (5) is a correspondence $k = k^*(\eta)$.

Proof. We use the fact that $F(k; \eta; n) = \gamma^n F(k; \eta; 0)$ to "guess" the trial solution: $\tilde{V}(\eta; n) = \gamma^n V(\eta)$ for some function V. Given the uniqueness of $\tilde{V}(\eta; n)$, this trial solution must be correct if it solves the DPE. Since the equation of motion of η is independent of n, we can substitute the trial solution into equation (5) to obtain an equivalent DPE

$$\gamma^n V(\eta_t) = \max_{k_t \in \{0,1\}} \gamma^n \{F(k; \eta_t; 0) + \beta\gamma^k V(\eta_{t+1})\}. \quad (6)$$

Dividing both sides by γ^n results in a DPE – and thus an optimal decision rule – which is independent of both n and t. ■

4 Choice of Technology

4.1 Myopic Case

Before analyzing the case for forward looking agents, we review JN's results for the case where agents base their current adoption decisions only on profits in the current period. In this case, agents solve the problem $\max\{\mu - \frac{1}{\eta}; \gamma(\mu - \frac{1}{h(\eta)})\}$, which uses the definition $\mu \equiv \frac{\gamma-1}{\gamma} > 0$ in equation (1), and the definition of $h(\eta)$. The first term in the maximand equals profits if the agent sticks with the current technology, and the second equals profits if the agent upgrades to the next generation of technology. (We ignore the factor γ^n, which affects profits under both alternatives, but not the adoption decision.) The agent sticks with the current technology if and only if the current precision satisfies the inequality

$$z(\eta) \equiv \frac{\alpha\gamma + \gamma\sigma_\epsilon^2\eta - 1}{\eta} - \mu(\gamma - 1) \geq 0:$$

The function $z(\eta)$ gives the increased profits, in the current period, resulting from not upgrading. In other words, $z(\eta)$ is the opportunity cost of adoption. The slope of $z(\eta)$ has the same sign as $1 - \alpha\gamma$. If there exists a positive root of $z(\eta) = 0$, it is unique. Denote this root (when it exists) as $\eta^c \equiv \frac{1-\alpha\gamma}{\gamma\sigma_\epsilon^2 - \mu(\gamma-1)}$. The agent is indifferent between upgrading and sticking if and only if $\eta = \eta^c$, i.e. when the opportunity cost of adoption is zero.

	$\alpha\gamma > 1$	$\alpha\gamma < 1$
$\sigma_\epsilon^2 > \frac{\mu(\gamma-1)}{\gamma}$	stagnation; never upgrade $\forall\eta^c$	(possible) overtaking; upgrade if $\eta < \eta^c$
$\sigma_\epsilon^2 < \frac{\mu(\gamma-1)}{\gamma}$	standard case; upgrade if $\eta > \eta^c$	continual upgrading $\forall\eta^c$

Table 1: The Myopic Model

Table 1 summarizes the relation between the parameter values and the optimal decision. In entries along the diagonal, it is optimal either never to upgrade or to upgrade in every period, regardless of the value of η. In these situations, η^c does not exist. In the lower left entry of Table 1, agents with low precision stick with the current technology until they learn to use it sufficiently well (until $\eta \geq \eta^c$), at which time they upgrade. We refer to this as the "standard" case.

In the upper right entry, it is optimal to upgrade only if the agent has low precision. An agent who is relatively unfamiliar with the current technology (i.e., has low precision $\eta < \eta^c$) upgrades, whereas the agent who knows how to use the current technology well (i.e. has high precision $\eta > \eta^c$) sticks with it. In this situation, the agent with lower initial precision (and thus, lower initial profits) may eventually obtain higher profits: she continues to upgrade her technology even though she never becomes expert at using it. In that sense, she overtakes the agent with high initial precision.

In order to guarantee that overtaking occurs, we need the following additional restriction. Define η_s as the (unique) positive steady state to equation (4).

Assumption 2 $\eta^c > \eta_s$.

The following lemma summarizes the overtaking result in JN.

Assumption 3 Lemma 1 (Overtaking) When $\alpha\gamma < 1$, $\sigma_2^2 > \frac{\mu(\gamma-1)}{\gamma}$, and Assumption 2 holds, an agent with initial precision $\eta < \eta^c$ eventually earns higher profits than an agent with initial precision $\eta > \eta^c$.

Proof. The agent with initial precision $\eta > \eta^c$ never upgrades, so $\eta_t \to \infty$ and her profits converge to $\gamma^{n_0}\mu$, where n_0 is the initial grade of technology. The agent with initial precision $\eta < \eta^c$ continues to upgrade in every period so $n_t \to \infty$ and $\eta_t \to \eta_s$. Thus, her profits approach $+\infty$ provided that $\mu - \frac{1}{h(\eta_s)} > 0$. Suppose to the contrary that $\mu - \frac{1}{h(\eta_s)} \leq 0$. In that case, $\gamma(\mu - \frac{1}{h(\eta_s)}) \leq \mu - \frac{1}{h(\eta_s)} < \mu - \frac{1}{\eta_s}$ (since $h(\eta_s) < \eta_s$), so it is not optimal to upgrade at η_s, contradicting the assumptions of the lemma. ■

If Assumption 2 did not hold, all agents would eventually cease to upgrade, and overtaking might not occur. Hereafter, when discussing the case of overtaking, we maintain Assumption 2.

4.2 General Case

This section generalizes the results from the myopic setting. All of the four possibilities described in Table 1 remain when β is positive. Thus, the possibility that overtaking occurs does not rely on the assumption that agents are myopic. However, if agents are sufficiently patient, overtaking cannot occur. We also show that a positive value of β never decreases, and typically increases the set of precision levels at which upgrading is optimal. In this sense, a forward looking agent upgrades more frequently than a myopic agent.

First we define overtaking in the general setting. Overtaking requires that there is an interval of η over which the agent is willing to upgrade. Moreover, if the initial precision lies in this interval, the equilibrium technology sequence is unbounded: $\lim_{t\to\infty} n_t = \infty$. There is also a critical value of η, which we denote $\bar{\eta}$, above which the agent never upgrades. Thus, if one agent begins with precision in the interval for which upgrading continues, and a second agent begins with $\eta > \bar{\eta}$, the first agent eventually uses higher grade technology and receives higher profits in every period, regardless of their initial technologies (their initial values of n).

The next two theorems analyze the first row of Table 1 when $\beta > 0$. Theorem 1 shows that overtaking is a generic possibility. Theorem 2 shows that a sufficiently large value of β eliminates the possibility of overtaking. If β is large it is optimal to sometimes upgrade, under configurations of parameter values for which upgrading is never optimal when $\beta = 0$ (i.e. in the upper left entry in Table 1).

Theorem 1 *If Assumptions 1 and 2 hold and $\alpha\gamma < 1$, $\sigma_2^2 > \frac{\mu(\gamma-1)}{\gamma}$ (so that overtaking occurs when $\beta = 0$), overtaking can occur for small positive values of β.*

Proof. We show that for sufficiently small but positive values of β; it is optimal to upgrade in every period when η is small, and it is optimal never to upgrade when η is large. Using equation (6) and the definition of $z(\eta)$, it is optimal not to upgrade if

$$z(\eta) > \beta[\gamma V(h(\eta)+\nu) - V(\eta+\nu)]: \tag{7}$$

$V(\eta)$ is nondecreasing and $V(\frac{1}{\mu}) > 0$; since the strategy of never upgrading in the future gives a stream of positive payoffs when $\eta > \frac{1}{\mu}$. Thus, for $\eta \geq \frac{1}{\mu}$, the right side of equation (7) is bounded above by $\beta\gamma V(h(\eta)+\nu)$. For all η, $\beta\gamma V(h(\eta)+\nu)$ is bounded above by $\frac{\beta\gamma\mu}{1-\beta\gamma}$(which equals the present value of the payoff if a new technology is adopted in every period and the precision instantly becomes infinite). Define η^* as the unique positive solution to $z(\eta) = \frac{\beta\gamma\mu}{1-\beta\gamma}$. Given the assumed parameter restrictions, η^* exists for β sufficiently small but positive. Thus, equation (7) is satisfied, and it is optimal not to upgrade for $\eta \geq \hat{\eta} \equiv \max\{\eta^*; \frac{1}{\mu}\}$.

It is optimal to upgrade if the inequality in equation (7) is reversed. The right side of equation (7) is approximately 0 for small β; the left side is independent of β and is strictly negative for η in the neighborhood of η_s (since $\eta_s < \eta^c$). Therefore, for sufficiently small β there exists a critical value of η greater than η_s, below which it is optimal to upgrade. If the initial value of η is below this critical value, the agent upgrades in every period. Since $\mu - \frac{1}{h(\eta_s)} > 0$ by lemma 1, overtaking occurs ■

Although the possibility of overtaking is generic, it never occurs if agents are sufficiently patient.

Theorem 2 Suppose Assumption 1 holds.

1. If $\alpha\gamma < 1$, $\sigma_\varepsilon^2 > \frac{\mu(\gamma-1)}{\gamma}$, and Assumption 2 holds (so that overtaking occurs when $\beta = 0$), there exists $\beta^* < \frac{1}{\gamma}$ such that for all $\beta \geq \beta^*$; overtaking cannot occur.

2. If $\alpha\gamma > 1$, $\sigma_\varepsilon^2 > \frac{\mu(\gamma-1)}{\gamma}$ (so that it is never optimal to upgrade when $\beta = 0$), and in addition $\mu - \frac{1}{h(\eta_s)} > 0$;[3] it is sometimes optimal to upgrade when $\beta \geq \beta^*$.

[3]When η^c does not exist, we obviously cannot invoke Assumption 2. We therefore impose this inequality directly.

Proof. 1. Overtaking requires that agents with sufficiently high precision never upgrade. We show that never upgrading in the future cannot be an optimal policy when β is large. Define $\pi_s \equiv \mu - \frac{1}{h(\theta_s)}$, which is positive by lemma 1. Therefore the value of the optimal program at θ_s is $V(\theta_s) \geq \frac{\pi_s}{1-\beta\gamma}$. The payoff from never upgrading is bounded above by $\frac{\mu}{1-\beta}$. Monotonicity of $V(\theta)$ implies that it is not optimal to stick with the current technology forever if $\frac{\pi_s}{1-\beta\gamma} > \frac{\mu}{1-\beta}$, i.e. if $\beta > \frac{\mu - \pi_s}{\gamma\mu - \pi_s} \equiv \beta^*$. Since $\gamma > 1$; $\beta^* < \frac{1}{\gamma}$. Thus there exists a range of parameter values that satisfy Assumptions 1 and 2 and $\alpha\gamma < 1$, $\sigma_2^2 > \frac{\mu(\gamma-1)}{\gamma}$, for which overtaking cannot occur.

2. The proof of part 2 uses the same argument to show that never upgrading is not optimal when β is sufficiently large. ■

We next show how forward-looking behavior changes the set of θ at which upgrading is optimal. We define $\theta^{c\beta}$ as a value of θ at which the agent with discount rate β is indifferent between sticking with the current technology and upgrading. That is, $\theta^{c\beta}$ satisfies

$$z(\theta) = \beta[\gamma V(h(\theta) + \nu) - V(\theta + \nu)] \tag{8}$$

(so $\theta^{c0} = \theta^c$). As with the static case, $\theta^{c\beta}$ may not exist, in which case the agent either upgrades in every period, or never upgrades. Unlike the static case, we have not shown that $\theta^{c\beta}$ is unique. When we refer to $\theta^{c\beta}$ we always mean any value of θ that satisfies equation (8). We show that $z(\theta^{c\beta}) > 0$ for $\beta > 0$. This inequality means that at a level of precision where the agent is indifferent between upgrading and sticking, upgrading reduces profits in the current period. We first state two facts which we use to prove this result.

Lemma 2 Define the function $\chi(\theta; w) \equiv (\gamma - 1)\mu + \frac{1}{\theta + w} - \frac{\gamma}{h(\theta)+w}$. $\chi(\theta; w)$ is an increasing function of w.

Proof. Differentiate the function χ and use the restriction that $\theta > h(\theta)$. ■

Lemma 3 $h(\theta) + \nu > h(\theta + \nu)$.

Proof. $h(\theta + \nu) < h(\theta) + h'(\theta)\nu < h(\theta) + \nu$ where the first inequality follows from concavity and the second from the restriction $\theta > h(\theta)$ which implies $h'(\theta) < 1$. ■

Proposition 3 *For $\beta > 0$, at a level of precision where the agent is indifferent between upgrading and sticking, upgrading causes losses in the current period: $z(\eta^{c\beta}) > 0$.*

Proof. Suppose to the contrary that

$$z(\eta^{c\beta}) \leq 0 \tag{9}$$

We derive a contradiction for the two interesting cases.[4] Case 1: it is optimal to upgrade at $\eta^{c\beta} - \epsilon$ for small positive ϵ and it is optimal to stick with the current technology for T periods at $\eta^{c\beta} + \epsilon$. (We allow the possibility that $T = \infty$, a necessary condition for overtaking.) Case 2: It is optimal to stick at $\eta^{c\beta} - \epsilon$ for small positive ϵ and it is optimal to upgrade at $\eta^{c\beta} + \epsilon$. (Case 2 corresponds to the second row of Table 1.)

Case 1. Choose $\eta = \eta^{c\beta} + \epsilon$, so that the optimal policy yields the payoff

$$V(\eta) = \sum_{t=0}^{T-1} \beta^t \left(1 - \frac{1}{\eta + t\gamma}\right) + \beta^T V(h(\eta + T\gamma))$$

where T (possibly infinite) is the optimal time of the next upgrade. Consider the deviation of moving forward the time of the next upgrade, e.g. upgrading at time 0 rather than time T. The payoff corresponding to this deviation is $D(\eta)$

$$D(\eta) = \sum_{t=0}^{T-1} \beta^t \left(1 - \frac{1}{h(\eta) + t\gamma}\right) + \beta^T V(h(\eta) + T\gamma).$$

Using these expressions, we have

$$D(\eta) - V(\eta) = \left[\sum_{t=0}^{T-1} \beta^t \chi(\eta; t\gamma)\right] + \beta^T \{V(h(\eta) + T\gamma) - V(h(\eta + T\gamma))\}.$$

Evaluate this difference at $\eta = \eta^{c\beta}$, where $\chi(\eta; 0) = -z(\eta) \geq 0$ by equation (9). By lemma 2, $\chi(\eta; t\gamma) > 0$ for $t > 0$, so the term in the square brackets is positive. By lemma 3 and monotonicity of V, the term in the curly brackets is positive. Therefore $D(\eta^{c\beta}) - V(\eta^{c\beta}) > 0$, which contradicts optimality.

---

[4] We ignore the unlikely possibility that the agent prefers to upgrade (or prefers to stick) for both $\eta^{c\beta} \pm \epsilon$, ϵ small. Even if this situation could arise, it is plausible that a perturbation of parameters would eliminate it.

Case 2. Choose $\eta \leq \eta^{c\beta}$ with $\eta + \nu > \eta^{c\beta}$. The optimal policy at such a value of η is to wait until the next period to upgrade, which leads to the payoff

$$V(\eta) = \left(\mu - \frac{1}{\eta}\right) + \beta\gamma\left(\mu - \frac{1}{h(\eta+\nu)}\right) + \beta^2\gamma V(h(\eta+\nu)+\nu).$$

Consider the deviation of upgrading in the current period rather than in the next one. We again denote the value of this deviation as $D(\eta)$:

$$D(\eta) = \gamma\left(\mu - \frac{1}{h(\eta)}\right) + \beta\gamma\left(\mu - \frac{1}{h(\eta)+\nu}\right) + \beta^2\gamma V(h(\eta)+2\nu).$$

The difference in the payoff is

$$\begin{aligned} D(\eta) - V(\eta) \;=\;& -z(\eta) + \beta\gamma\left[\frac{1}{h(\eta+\nu)} - \frac{1}{h(\eta)+\nu}\right] \\ &+ \beta^2\gamma\left\{V(h(\eta)+2\nu) - V(h(\eta+\nu)+\nu)\right\}. \end{aligned}$$

Evaluate this difference at $\eta = \eta^{c\beta}$. The first term on the right side is non-negative by equation (9), the second term (square brackets) is positive by lemma 3, and the third term (curly brackets) is positive by lemma 3 and the monotonicity of the value function. Consequently, $D(\eta) - V(\eta) > 0$, which contradicts optimality. ■

Proposition 3 demonstrates the trade-off the agent faces in the choice of technology under learning-by-doing. Forward-looking agents upgrade to the new technology because the future benefit from the new technology exceeds the short-term cost from discarding the familiar old technology. In other words, forward looking agents upgrade when the current payoff from the new technology is strictly less than the current payoff from the old technology, whereas myopic agents upgrade only when the current payoff from the new technology is at least as great as the current payoff from the old technology.

Proposition 3 enables us to compare the critical values $\eta^{c\beta}$ and η^c. In order to allow for the possibility that $\eta^{c\beta}$ is not unique, we define $\bar{\eta}^{c\beta} = \max\{\eta^{c\beta}\}$ and $\underline{\eta}^{c\beta} = \min\{\eta^{c\beta}\}$. We have

Corollary 1 Suppose $\eta^{c\beta}$ and η^c exist. Then $\bar{\eta}^{c\beta} < \eta^c$ for $\alpha\gamma > 1$; and $\underline{\eta}^{c\beta} > \eta^c$ for $\alpha\gamma < 1$.

Proof. By inspection, $z(\theta)$ is monotonic, and the derivative $\frac{dz}{d\theta}$ has the same sign as $1 - \alpha\gamma$. From Proposition 3, $z(\theta^{c\beta}) > 0 = z(\theta^{c})$. Hence, when $\frac{dz}{d\theta} > 0, \theta^{c\beta} > \theta^{c}$, implying $\underline{\theta}^{c\beta} > \theta^{c}$ for $\alpha\gamma < 1$. When $\frac{dz}{d\theta} < 0, \theta^{c\beta} < \theta^{c}$, implying $\tilde{\theta}^{c\beta} < \theta^{c}$ for $\alpha\gamma > 1$. ■

Figure 1, which shows the graph of $z(\theta)$ for the two cases $\alpha\gamma < 1$ (solid curve) and $\alpha\gamma > 1$ (dashed curve), illustrates the corollary. We use this result to show how a positive value of β affects the decision to upgrade. Define the "upgrade set" $\Theta^{\beta} = \{\theta : k(\theta) = 1\}$, the set of θ for which it is optimal to upgrade, given β. Table 1 implicitly defines Θ^{0} (the upgrade set for $\beta = 0$) under different configurations of parameter values. The following theorem compares the upgrade sets for $\beta = 0$ and for $0 < \beta < \frac{1}{\gamma}$ under these four configurations of parameter values.

Theorem 3 For $0 < \beta < \frac{1}{\gamma}$, $\Theta^{0} \subseteq \Theta^{\beta}$.

Proof. It is convenient to prove the claim for separate cases in Table 1. That is,

(i) If, for $\beta = 0$, there is either stagnation ($\alpha\gamma > 1$ and $\sigma_{\epsilon}^{2} > \frac{\mu(\gamma-1)}{\gamma}$) or continual upgrading ($\alpha\gamma < 1$ and $\sigma_{\epsilon}^{2} < \frac{\mu(\gamma-1)}{\gamma}$), then $\Theta^{0} \subseteq \Theta^{\beta}$.

(ii) If the "standard case" occurs when $\beta = 0$ ($\alpha\gamma > 1$ and $\sigma_{\epsilon}^{2} < \frac{\mu(\gamma-1)}{\gamma}$), then $\Theta^{0} \subset \Theta^{\beta}$.

(iii) If overtaking is possible when $\beta = 0$ ($\alpha\gamma < 1$, and $\sigma_{\epsilon}^{2} > \frac{\mu(\gamma-1)}{\gamma}$), then $\Theta^{0} \subset \Theta^{\beta}$.

We take these cases in turn.

(i) Under stagnation, $\Theta^{0} = \emptyset \subseteq \Theta^{\beta}$. (From Theorem 2, Θ^{β} may be nonempty, in which case $\Theta^{0} \subset \Theta^{\beta}$.) Under continual overtaking, $\Theta^{0} = \Re^{+}$, and it is straightforward to show that $\Theta^{\beta} = \Re^{+}$.

(ii) In this case, $\Theta^{0} = \{\theta : \theta > \theta^{c}\}$. If $\theta^{c\beta}$ exists, then it must be the case that $\Theta^{\beta} \supseteq \{\theta : \theta > \tilde{\theta}^{c\beta}\}$. If this relation did not hold, then for sufficiently large θ, it is optimal never to upgrade. However, using the inequality $\sigma_{\epsilon}^{2} < \frac{\mu(\gamma-1)}{\gamma}$ we can show that for sufficiently large θ the payoff of upgrading once and then never subsequently upgrading is greater than the payoff of never upgrading. Since

$\tilde{\eta}^{c\beta} < \eta^c$ from Corollary 1, we obtain $\Delta^\beta \supseteq \{\eta : \eta > \tilde{\eta}^{c\beta}\} \supset \Delta^0$. If $\eta^{c\beta}$ does not exist, it is optimal to upgrade for all η, so $\Delta^\beta = \mathbb{R}^+$.

(iii) In this case, $\Delta^0 = \{\eta : \eta \leq \eta^c\}$. If $\eta^{c\beta}$ exists, then from Corollary 1, $\underline{\eta}^{c\beta} > \eta^c$. We need to show that $\Delta^\beta \supseteq \{\eta : \eta < \underline{\eta}^{c\beta}\}$. (This relationship implies that for $\beta > 0$ it is strictly better to upgrade at $\eta = \eta^c$.) Suppose, to the contrary, that for $\eta < \underline{\eta}^{c\beta}$ it is optimal not to upgrade. Then at $\eta = \eta^c$ it is optimal to stick with the current technology for $T \geq 1$ periods, where T is the smallest integer that satisfies $\eta = \eta^c + T\nu \geq \eta^{c\beta}$. At time T time it is optimal to upgrade. Consider the deviation of upgrading in the current period (when $\eta = \eta^c$) rather than waiting T periods. The additional profits resulting from this deviation, rather than following the optimal program, are

$$\left[\sum_{t=0}^{T-1} \beta^t \chi(\eta^c; t\nu)\right] + \beta^T \{V(h(\eta^c) + T\nu) - V(h(\eta^c + T\nu))\}.$$

The first term (square brackets) is positive using the definition of η^c and lemma 2, and the second term (curly brackets) is positive by lemma 3 and monotonicity of $V(\cdot)$. Consequently, it must be optimal to upgrade when $\eta = \eta^c$ and $\beta > 0$. Therefore $\Delta^\beta \supseteq \{\eta : \eta < \underline{\eta}^{c\beta}\} \supset \{\eta : \eta < \eta^c\} = \Delta^0$.

If $\eta^{c\beta}$ does not exist, it is optimal to upgrade for all η, so $\Delta^\beta = \mathbb{R}^+$. ■

Theorem 3 means that forward-looking agents are "more likely" to upgrade than myopic agents. For example, if overtaking occurs in the myopic setting, the introduction of a positive discount factor reduces (and according to Theorem 2 may eliminate) the values of η above which further upgrading never occurs. In addition, if $\sigma_\epsilon^2 < \frac{\mu(\gamma-1)}{\gamma}$ (the second row in Table 1) so that overtaking does not occur when $\beta = 0$, then overtaking cannot occur when $\beta > 0$. Finally, if $\sigma_\epsilon^2 > \frac{\mu(\gamma-1)}{\gamma}$, $\alpha\gamma > 1$ (the upper left entry in Table 1) myopic agents would never upgrade. For these parameter values and $\beta > 0$; agents might upgrade when η is sufficiently large. In this case, the introduction of a positive discount factor transforms the "stagnation" scenario to the "standard" scenario, in which agents wait until they are sufficiently familiar with the current technology before upgrading.

Theorem 3 compares the upgrade sets under a myopic and a forward looking agent. The next theorem compares upgrade sets for small values of β. To emphasize the dependence of the value function on β, we replace $V(\eta)$ with $V(\eta;\beta)$.

Theorem 4 *For small β the upgrade set is monotone in β: $\Delta^{\beta} \subseteq \Delta^{\beta'}$ for $\beta' > \beta$ (with β' and β small).*

Proof. Define the function

$$G(\eta;\beta) \equiv z(\eta) - \beta\left[\gamma V(h(\eta)+\nu;\beta) - V(\eta+\nu;\beta)\right].$$

$G(\eta;\beta)$ is the gain from sticking with the current technology, and $\eta^{c\beta}$ satisfies $G(\eta^{c\beta};\beta) = 0$. From equation (7) it is optimal to upgrade if and only if $G(\eta;\beta) \leq 0$; i.e. $\Delta^{\beta} = \{\eta : G(\eta;\beta) \leq 0\}$. The implicit function theorem can be invoked to yield

$$\frac{d\eta^{c\beta}}{d\beta} = -\frac{G_{\beta}(\eta^{c\beta};\beta)}{G_{\eta}(\eta^{c\beta};\beta)}.$$

Continuity of $G(\eta;\beta)$ implies that $\Delta^{\beta} \subseteq \Delta^{\beta'}$ iff $G_{\beta}(\eta^{c\beta};\beta) < 0$. (If $G_{\eta}(\eta^{c\beta};\beta) < 0$, it is optimal to upgrade for $\eta = \eta^{c\beta} + \varepsilon$, so a decrease in $\eta^{c\beta}$ enlarges the upgrade set. If $G_{\eta}(\eta^{c\beta};\beta) > 0$, it is optimal to upgrade for $\eta = \eta^{c\beta} - \varepsilon$, so an increase in $\eta^{c\beta}$ enlarges the upgrade set.) From the definition of $G(\cdot)$ we have

$$\begin{aligned}\frac{\partial G}{\partial \beta} &= -\left[\gamma V(h(\eta)+\nu;\beta) - V(\eta+\nu;\beta)\right] \\ &\quad -\beta\left(\gamma\frac{\partial V(h(\eta)+\nu;\beta)}{\partial \beta} - \frac{\partial V(\eta+\nu;\beta)}{\partial \beta}\right)\end{aligned}$$

From Proposition 3, the term in square brackets is positive (since it equals $z(\eta^{c\beta}) > 0$). The derivative of $V(\eta+\nu;\beta)$ with respect to β is positive and bounded above by $\frac{\gamma}{(1-\beta\gamma)^2}$ since the value function is bounded above by $\frac{1}{1-\beta\gamma}$ as shown in Proposition 1. Consequently the term on the second line can be made arbitrarily small by choosing small β. It follows that for small β the upgrade set is monotone in β. ■

Theorem 4 implies that as the agent becomes more patient, she upgrades to the new technology for a larger set of precision levels. The combined results of Theorems 1-4 can be related to the role of capital markets and technology choice at di®erent stages of development.[5] At an early stage of economic development, an economy may have no ¯nancial markets, causing the cost of capital and the discount rate to be high. In this case overtaking occurs for some initial conditions and parameter values. As the economy develops, ¯nancial markets also develop, leading to a lower discount rate and causing ¯rms to become more willing to upgrade. Since overtaking is less likely for patient agents, convergence is more likely to occur among developed economies, and certainly occurs if the discount rate becomes su±ciently small.

5 Conclusion

When skills are only partly transferrable across generations of technology, being more expert at using an existing technology may make it easier adopt a higher grade. However, greater skill at using the existing technology also leads to a higher opportunity cost of upgrading. The agent who is skilled at using the existing technology may decide not to upgrade. An agent who is less skilled has a lower opportunity cost and may upgrade, even though she cannot use the new technology as pro¯tably as the ¯rst agent. The less skilled agent may continue to upgrade to increasingly sophisticated technologies, even though she never becomes expert at using any of them. She eventually achieves higher pro¯ts than the more skilled agent.

This kind of overtaking can occur even when agents are forward looking, as in Parente's (1994) model. However, overtaking never occurs if agents are su±ciently patient. Previous papers emphasized the characteristics of technology and learning that lead to the possibility of overtaking. We have emphasized the need for a su±ciently low discount factor in order to obtain overtaking.

We also showed that when the myopic agent's upgrade decision depends non-trivially on her skill level, a forward looking agent decides to upgrade for a larger set of skill levels. At least for small discount factors, the upgrade set is nondecreasing in the discount factor. In this sense, forward looking agents are more

[5]Parente (1994) made a similar point based on a simulation result that the availability of capital market a®ects the develpment path via the technology choice.

likely to upgrade, and they upgrade more frequently.

Overtaking occurs because the skilled agent, who can use the new technology more profitably, also has a higher opportunity cost of upgrading, relative to the unskilled agent. In the model we analyzed, the higher opportunity cost is the result of learning-by-doing. However, other actions taken in the past might also give rise to overtaking. For example, a producer who uses the current technology may find it cheaper to upgrade to a newer technology, relative to a producer who is not using the current technology. However, the first producer also has less incentive to adopt the newer technology, because of the alternative to continue using the technology which has already been purchased.

References

[1] Aghion, P., and P. Howitt, Endogenous Growth Theory, (1998), MIT Press, Cambridge.

[2] Barro, R. J., and X. Sala-i-Martin, Economic Growth, (1995), McGraw-Hill, New York.

[3] Brezis, E. S., P.R. Krugman, and D. Tsiddon, Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership, American Economic Review, 83, (1993), 1211-1219.

[4] Budd, C, C Harris and J Vickers, A Model of the Evolution of Duopoly: Does the Asymmetry Between Firms Tend to Increase of Decrease? Review of Economic Studies, 60 (1993) 543 - 73.

[5] Chari, V. V., and H. Hopenhayn, Vintage Human Capital, Growth, and the Di®usion of New Technology, Journal of Political Economy, 99, (1991), 1142-1165.

[6] DeGroot, M. H., Optimal Statistical Decision, (1970), McGraw-Hill, New York.

[7] Jovanovic, B., and Y. Nyarko, Learning-by-doing and the Choice of Technology, 64, Econometrica, (1996), 1299-1310.

[8] Krusell, P., and J.-V. Rios-Rull, Vested Interests in a Positive Theory of Stagnation and Growth, Review of Economic Studies, 63, (1996), 301-329.

[9] Lucas, R. E., Making a Miracle, Econometrica, 61, (1993), 251-272.

[10] Parente, S., Technology Adoption, Learning-by-doing and Economic Growth, Journal of Economic Theory, 63, (1994), 349-369.

[11] Motta, M., J.F. Thisse, and A. Cabrales, On the Persistence of Leadership or Leapfrogging in International Trade International Economic Review, 38, (1997), 809 - 824.

[12] Stokey, N. L., Learning-by-doing and the Introduction of New Goods, Journal of Political Economy, 96, (1988), 701-717.

[13] Stokey, N. L., and R. E. Lucas, Jr., Recursive Methods in Economic Dynamics, (1989), Harvard University Press, Cambridge and London.

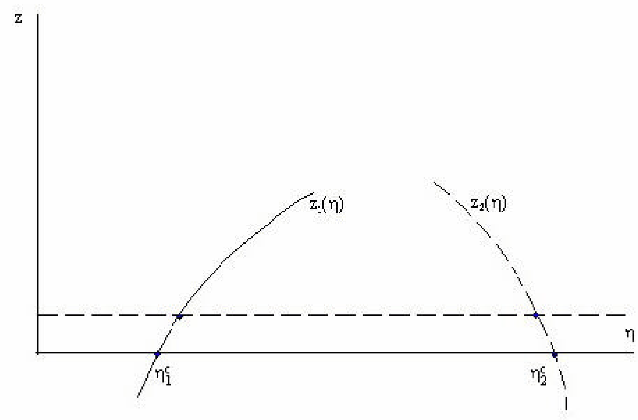

Figure 1