Department of Economics
University of Southampton
Southampton SO17 1BJ
UK

**Discussion Papers in
Economics and Econometrics**

2000

**This paper is available on our website
http://www/soton.ac.uk/~econweb/dp/dp00.html**

Technology Adoption with Finite Horizons

Xavier Mateos-Planas
Department of Economics
University of Southampton
Highfield SO17 1BJ
Southampton, UK
E-Mail: fxmp@soton.ac.uk

Version 1.0
10 November 2000

Abstract

This paper analyzes the optimal sequence of technology upgrades by a firm that lives for a finite period of time. Other characteristics of the environment are the existence of technology-specific learning-by-doing, technology growth, and sunk costs. A finite planning horizon implies that the technology adoption problem is non-stationary and the frequency of adoptions changes over time. This paper provides results for the computation of the optimal plan and explores numerically the life-cycle pattern of technology switches.

Keywords: learning-by-doing, technology adoption.

1 Introduction

The adoption of technologies is an engine of economic progress. Even in the more developed economies, the resources devoted to technology adoption are substantial relative to those to technology-creating activities or research [see Jovanovic (1995)]. Different patterns of technology adoption are also invoked as part of the explanation for observed disparities in economic performances. These include economic inequalities across countries, as well as across workers [Doms et al. (1997), Bartel and Litchenberg (1987), Parente and Prescott (1994)]. Broadly viewed, decisions in many spheres of life involve elements that resemble the choice of adopting new technologies. Examples include a government considering to push ahead with policy reforms, or the decision to change a job or career by a worker. Therefore, the study of the factors that shape the patterns of the technology adoption choices is important to understand a variety of interesting choice problems.

The aim of this paper is to contribute to the analysis of technology adoption decisions in dynamic contexts. The specific objective of this paper is to solve and characterize the pattern of technology switches when there is a finite time-horizon for the agent. This is a pervasive characteristic of environments where technology adoption-like decisions are made. In many labor markets workers confront a strictly finite work-life dictated by the retirement age. Fixed-term labor contracts have a predetermined termination data. In some countries, public utilities are managed by private firms over a predetermined period before they revert back to the government. Heads of government can be reelected only a finite number of times. Patents typically guarantee a protection for a finite period. The goal of the present paper is to explore the implications of this upper bound for the pattern of technology switches.

The analysis of the finite-horizon case may also prove useful to analyze models where the time horizon is infinite but the discount rate may change over time. The sequence of technology adoptions can be regarded as a succession of finite-horizon problems. This feature is characteristic, for example, of dynamic general equilibrium models where the interest rate changes over the transition. Time-changing discount rates are also a feature of life-cycle models where agents of different ages face different survival probabilities.

The paper analyzes the technology switching problem of a single agent with a finite horizon in continuous time. There is a process of continuously improving technologies which

provides the drive for switching technologies. There are costs to switching technologies though. On one hand, there is direct sunk cost associated with the upfront investment needed to implement the technology to be adopted. On the other hand, learning-by-doing on the current technology has to be foregone on adopting a new one. In this model, the process governing the emergence of new technologies and learning-by-doing are deterministic. The key ingredient of the model is the choice by finitely-lived agents on technology adoptions in the presence of exogenous embodied technological change and technology-specific learning-by-doing. This paper analyzes the choice of multiple technology adoption as a non-stationary dynamic programming problem where both the number of adoptions as well as their timing are the choice variables. The solution allows to investigate the properties of the pattern of technology adoptions under these circumstances.

The outcomes of this paper are as follows. The analysis provides an algorithm for the solution of the problem which exploits the analytical features of the model. An optimal plan may include technologies that are learned along with technologies that are replaced before learning occurs. In those cases, it is shown that the adoptions where learning occurs must necessarily take place first. The pattern of adoptions will in general be uneven. The qualitative and quantitative implications depend on the scope for and speed of learning-by-doing, the costs of adoption, the discount rate, and the rate of technological progress. The paper illustrates through numerical analysis the effects of these factors.

This paper relates to a body of literature that analyzes the replacement/adoption of technologies in dynamic settings, and that includes Zeckhauser (1968), Stokey (1991), Jovanovic and Nyarko (1996), Parente (1994, 2000), Cooley et al. (1997), Greenwood and Yorukoglu (1996), Jovanovic and Rob (1998), Klenow (1998), Yorukoglu (1998), and Mateos-Planas (forthcoming). These papers study infinite-horizon problems and the optimal choices are typically characterized by a constant pattern of technology adoption. In the present paper, instead, the planning horizon is finite which implies that the policy function is non-stationary. A simplifying feature in the present paper is that learning-by-doing is technology specific. Other papers, like Stokey (1991), Jovanovic and Nyarko (1996), and Parente (1994, 2000), accommodate the transfer of knowledge across different technologies. Greenwood and Yorukoglu (1996) and Cooley et al. (1997) study a discrete-time model and the solution is approximated numerically. The present paper is in continuous-time and the approach is more analytical.

There is also a body of literature on optimal investment and technology in stochastic models. It includes Kamien and Schwartz (1972), Jensen (1982), Balcer and Lippman (1984), McDonald and Siegel (1986), and Dixit and Pyndyck (1994). This literature emphasizes the role of uncertainty and market structure. The present paper studies a deterministic model for a competitive agent.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 describes the decision problem and outlines the approach taken for its solution. Section 4 contains the results that characterize some properties of the optimal choice and allow its explicit computation. Section 5 demonstrates the properties of optimal plans. Section 6 concludes the paper.

2   The model

The agent lives and produces output for a period of length T. The agent is assumed to operate a single firm over his productive life. When the agent dies the firm is dissolved and there is no market for discontinued firms.

The firm produces output using one machine. The flow of output of a firm at time t depends on the quality of the machine in use, and on the agent's technology-specific expertise. The quality of the machine is given by the technology embodied in it and I index technologies over the positive real line by a. Expertise in a technology is denoted by q. Output of a firm that operates a machine of quality a with technology-specific expertise q is

$$y = q \cdot a, \tag{1}$$

with and $a, q \in R_+$. At any instant of time, a firm may either switch to a more advanced technology or continue to use the present one. I call *technology adoption* the decision to operate a new technology by replacing the current machine with another of different quality.

The level of expertise on a technology evolves with its use as the result of learning-by-doing. Thus one can write q as a non-negative function $q(m)$, where m denotes the duration of use of the technology. This learning-by-doing is technology-specific. Thus if the firm decides to switch technologies, no part of the expertise in the previous technology can be carried over to the new one. This is a simplification with respect to Parente (1994)

4

and Jovanovic and Nyarko (1996).

The upper bound on the technologies that can be used by the firm at time t is denoted by $A(t)$. This frontier technology grows at a constant and exogenous rate γ over time. Switching to a technology a involves a cost to the firm of size $\pi \cdot a$ units of output. This fixed payment is meant to reflect the cost of the piece of capital that embodies the technology. This cost is a sunk cost. There is a perfect capital market where agents can borrow and lend at a constant interest rate r.

The agent maximizes the present life-time value of output- net of adoption costs- from the firm he operates. To this end, he decides which technology, among those available to him, to use at every instant over his productive life, $[0, T]$. The parameters that the agent takes as given are the learning curve $q(.)$, technological progress γ, the interest rate r, and the time horizon T. The shape of $q(.)$ will be specified later. A *feasible adoption plan* defines the set of choices available to the agent.

Definition-1 *Given parameters T and the path for technology $A(t)$, a feasible adoption plan is defined by:*

    i. An integer number, J, denoting the number of adoptions.

    ii. A sequence of real numbers $\{x_j\}$ for $j = 1, ..., J, J+1$ representing the dates at which each j-th adoption occurs, such that $0 \quad x_j < T$, $x_{j+1} > x_j$ for $j = 1, ..., J$, and $x_{J+1} = T$.

    iii. A path for the firm's technology $a(t)$ for $t \geq 0$ such that $a(t) \quad A(t)$, and $a(t)$ is constant for $t \in (x_j, x_{j+1})$ all $j = 1, ..., J$.

The technology in Eq. (1) implies that a feasible adoption plan generates a path of output $y(t)$ such that, for $j = 1, ..., J$,

$$y(t) = a(x_j)q(t - x_j), \quad t \in [x_j, x_{j+1}) \tag{2}$$

3 The Technology Adoption Problem

The problem of the firm consists of maximizing the present value of output net of adoption costs by choice of an appropriate feasible adoption plan. Let $V(\underline{x}, \overline{x})$ denote the optimal

value of the firm between the initial date $\underline{x}$ and a terminal date $\overline{x}$ divided by the initial level of technology. The maximization problem of the firm is then,

$$V(0,T) = \max_{\{J,x_j\}} \sum_{j=1}^{J} e^{-rx_j} a(x_j) W(x_{j+1} - x_j) \tag{3}$$

where

$$W(m) \equiv \int_0^m e^{-rt} q(t) dt - \pi, \tag{4}$$

and J and x_j's belong to the set of feasible adoption plans. Here $W(m)$ is the present value as of time 0 of output produced with a technology $a = 1$ over an interval of length m, minus the cost of adopting that technology. I will make the following assumption.

Assumption 1: $W(T) > 0$.

Assumption 2: $r - \gamma > 0$.

The first assumption simply means that activity has a non-negative value. The second assumption means that discounting must be sufficiently high. Under these assumptions it is straightforward to argue that an optimal adoption plan exists where the first adoption takes place at time 0, $x_1 = 0$, and the adopted technology is always the frontier so $a(x_j) = A(x_j)$ all j. This results follows from the assumption that neither the relative adoption costs, π, nor the speed of learning, $q(.)$, depend on the productivity of the technology to be adopted. Therefore, if an adoption occurs at time t, the technology adopted will be the frontier technology $a(t)$. Thus the agent's choice consists of deciding at every date $t \in [0, T]$ whether to keep on operating the current technology or switch to the frontier technology. Which technology is currently used influences the time at which the next technology is introduced but has no influence on the choice of which technology to adopt at that date.[1] Then the solution shows that optimal technology adoption results in a sequence of dates at which the firm switches to the frontier technology and stays there until the next upgrade. Assuming that $a(0) = 1$, this result allows us to rewrite the agent's problem as

$$V(x_1, T) = \max_{\{J,x_j\}} \sum_{j=1}^{J} e^{-(r-\gamma)x_j} W(x_{j+1} - x_j). \tag{5}$$

---

[1] Jovanovic and Rob (1998) also assume adoption of the frontier technology. Parente (1994), Jovanovic and Nyarko (1996), and Hendricks (1997) analyse adoption problems where the distance between the current and the new technology affect the costs of adoption.

A solution must specify the timing, x_j, and number, J, of adoptions. By increasing the frequency of adoptions, the firm is closer to the technology frontier more often. However, it has to pay adoption costs more often and reduces the benefits from learning. The optimal choice resolves this trade off. The choice of the number J is a novel feature of this analysis. As a benchmark for the results to come, when the horizon is infinite and $J \to \infty$ then a solution must consist of a sequence of equally spaced adoptions. The departure from this case will in general lead to a time-varying time span between consecutive adoptions.

It is useful to start by solving for the timing, taking an arbitrary J as given. The structure of this problem is recursive: the optimal decision rule mapping x_j into x_{j+1} for any $j = 1, 2, ..., J-1$ depends on optimal decision rules for future adoptions. Every adoption is chosen taking into account that subsequent adoptions will be decided optimally given the remaining time span. To be general, let $V(x, \overline{x}|k)$ denote the optimal value of a firm that lives between dates x and $\overline{x}$, conditional on the plan containing exactly k adoptions. Thus, if k is optimal, then $V(x, \overline{x}) = V(x, \overline{x}|k)$, where $V(., .)$ is as defined in Eq.(5) upon letting $x_1 = x$ and $T = \overline{x}$. Then one can write the problem recursively as follows:

$$V(x, \overline{x}|k) = \max_{x' \in [x, \overline{x}]} \left\{ W(x' - x) + e^{-(r-\gamma)(x'-x)} V(x', \overline{x}|k-1) \right\}. \tag{6}$$

The state for this choice is given by the current date, x, and the number, k, of adoptions contained in the plan that starts at this date. One can write the optimal choice as a policy function $m(.|.)$ that gives the optimal duration of use of the current adoption so $x' = x + m(x|k)$.

With these pieces of notation, the problem of the agent in equation (5) can be broken down into a sequence of problems as follows.

$$V(x_j, T|J - (j-1)) = \max_{x' \in [x_j, T]} \left\{ W(x' - x_j) + e^{-(r-\gamma)(x'-x_j)} V(x', T|J - j) \right\}, \tag{7}$$

for $j = 1, ..., J - 1$. With the convention that a plan involving zero adoptions carries zero value $V(., T|0) = 0$, it follows that $V(x_J, T|J - (J-1)) = W(x_{J+1} - x_J)$. Then the policy functions give the optimal sequence of tenures $m_j = m(x_j|J - (j-1))$ and $x_{j+1} = x_j + m_j$ for $j = 1, ..., J - 1$, and $m_J = x_{J+1} - x_J$.

7

Clearly this is a non-stationary dynamic programing problem for two reasons. First, the discount rate is changing over time (besides being affected by the choices). Second, the value functions depend on the order of the current technology j. For the given J, this problem could be solved backwards numerically by constructing grids for the current state. But this proves to be a highly inefficient procedure and provides no insight about the nature of the optimal plan. This paper will exploit the analytical properties of the problem to derive results that allow the computation of the exact optimal choices. Solving the problem involves to solve for $V(x, T; J - (j - 1))$ for different x and $j = 1, ..., J$. This recursion starts from the last-stage optimal choice, $V(x, T|1)$, and leads to the solution for the entire sequence $V(x, T|J)$.

Of course, for the given time span T, the arbitrary number of adoptions J may be inconsistent with an optimal choice. The second part of the problem is then to find the optimal J as the solution to

$$J = \arg \max_k \left\{ V(x_1, T|k) : k = 1, 2... \right\}, \tag{8}$$

The solution of the original problem in Eq. (5) is then $V(x_1, T) = V(x_1, T|J)$.

The approach of this paper to solving the problems defined in equations (7) and (8) is as follows. For given J, if a solution exists to Eq. (7) it must feature

$$V(x_j, T|J - (j - 1)) = W(x_{j+1} - x_j) + e^{-(r-\gamma)(x_{j+1} - x_j)} V(x_{j+1}, T|J - j) \tag{9}$$

for $j = 1, ..., J$. I will deal with situations where the value functions are differentiable and the solution can be characterized as a sequence that solves a first-order condition. Provided that the envelope theorem holds, the above Eq. (9) implies that the optimal interior choice of x_{j+1} in problem (7) must satisfy

$$W'(x_{j+1} - x_j) - e^{-(r-\gamma)(x_{j+1} - x_j)} \left[(r - \gamma) W(x_{j+2} - x_{j+1}) + W'(x_{j+2} - x_{j+1}) \right] = 0 \tag{10}$$

for $j = 1, ..., J - 1$. This expression has a clear interpretation in terms of the costs and benefit of delaying the date of the next adoption. This is a 2nd order difference equation in x_j with initial and terminal conditions $x_1 = 0$ and $x_{J+1} = T$, respectively. Similarly, it can be regarded as a 1st order difference equation in $m_j = x_{j+1} - x_j$ with $\sum^J m_j = T$.

Whether this condition is sufficient to characterize a solution, or gives the only solution will depend on the assumptions underlying process of technology-specific skill $q(.)$.

I will assume $q(.)$ is a non-decreasing function of time and has an upper bound. The first assumption rules out depreciation of skill with time of use. The second assumption implies bounded learning which is consistent with the empirical literature like Jovanovic and Nyarko (1995), Bahk and Gort (1993), Argote and Epple (1990), and Rapping (1965). One possible specification for the learning technology is the following continuous curve $q(t) = \delta + (1 - \delta) \exp(\mu t)$. Here δ represents the progress ratio, or the maximum factor increase in productivity that learning can produce. On its part, μ is a measure of the speed of learning. This learning curve has been used in Parente (1994). One problem with this specification is that, in general, the first-order condition is not sufficient for a maximum. In other words, more that one root x_{j+1} to Eq.(10) may exists, possibly implying a local minimum. This prevents the development of the approach in this paper that is based on solving the first-order condition. Henceforth another simpler process of learning will be considered. In particular, the following discrete-learning curve is assumed.

$$
q(t) = \begin{cases} 1 & \text{if } t < \mu \\ \delta & \text{otherwise} \end{cases} \tag{11}
$$

with $\delta > 1$. If the firm's experience in the use of its current technology is shorter than a period of length μ, its level of expertise in this technology is 1. Thereafter, its level of expertise in this technology increases to δ, which represents the progress ratio.

Even under this specification, the properties of the first-order condition in Eq. (10) do not rule out multiple local extrema. However, a method can be developed that allows us to deal with this circumstance. The first step is based on solving, separately, for plans where no technology is ever learned and plans where learning occurs in all technologies. These "restricted" plans are shown to have a solution that can be characterized by applying Eq. (10) for a given number of adoptions. This will be the result in Proposition 1.

Of course, within each class of plans, a given J may not be consistent with optimality in the sense that it is not possible to find a feasible sequence that satisfies the recursion in Eq. (10). Similarly, for a given initial date there may be different possible paths governed by the first-order condition in Eq. (10) that are be consistent with feasibility. Numerically

finding the number of adoptions may be costly. The result in Proposition 2 below allows us to determine exactly the optimal number of adoptions. To determine the optimal J, it is possible to partition the time interval into segments. Then initial dates on the real line can be mapped into the "restricted" optimal number of adoptions using this partition.

The two previous results characterize the restricted plans. The solution to the original plan in Eq. (5) may consist of one of the restricted plans or a combination of restricted adoption plans. In the latter case, and under a fairly mild assumption, the result in proposition 3 will show that adoptions where learning takes place must occur first. Results are provided that allow to identify conditions where only one class of restricted plan applies throughout or, otherwise, to narrow down the region of search.

4   Optimal Adoption of Technologies

With the specification of learning-by-doing in Eq. (11) above, one difficulty is that, in general, one has to account for the possibility that learning may not occur on some technologies that are adopted. Due to the discontinuity in the derivative of $W(.,.)$, there may be multiple local extrema at each stage of recursion in Eq. (7). Therefore, to characterize a solution it proves useful to consider two classes of restricted adoptions plans separately: plans that feature tenures shorter than μ only, which I call *S-plans*, and plans that feature tenures longer than μ only, *L-plans*. The restricted return functions, value functions and adoptions plans will be indexed by $v = L, S$ accordingly as $W^v(.,.)$, $V^v(.,.)$, m_j^v and x_j^v. In particular, given the definition in (4), the present value within a technology over an interval m can be written

$$W^v(m) = \begin{cases} \frac{1}{r}[1 - e^{-rm}] - \pi & \text{if } v = S \\ \frac{1}{r}[1 + (\delta - 1)e^{-r\mu} - \delta e^{-rm}] - \pi & \text{if } v = L \end{cases} \tag{12}$$

Section 4.1 characterizes these restricted plans. Section 4.2 then derives the (unrestricted) optimal adoption plan which may correspond to the optimal L-plan, the optimal S-plan, or a combination of L-plans and S-plans over different subperiods of time.

4.1 Characterization of restricted optimal plans

The optimal v-plan must satisfy a simple set of first order conditions. Notation is greatly simplified by defining,

$$\Gamma^v(m, m') \equiv I_1 - e^{\gamma m} \left[\frac{r - \gamma}{r} (I_2 - r\frac{\pi}{h}) + I_3 \frac{\gamma}{r} e^{-rm'} \right],$$

with

$$(I_1, I_2, I_3) = \begin{cases} (1, 1, 1) & \text{if } v = S \\ (\delta, 1 + (\delta - 1)e^{-r\mu}, \delta) & \text{if } v = L \end{cases} \tag{13}$$

The following proposition is proved in appendix A.

Proposition 1. *Suppose that x_j^v is the jth adoption in an optimal v-plan that ends at $\overline{x}$. Then the continuation optimal v-plan over $[x_j^v, \overline{x}]$, $\{x_j^v, x_{j+1}^v, ..., x_J^v\}$, is unique and must satisfy the sequence of first order conditions*

$$\Gamma^v(m_i^v, m_{i+1}^v) = 0, \quad \text{for } j = J - 1, ..., j, \tag{14}$$

and

$$\sum_{i=j}^{J} m_i^v = \overline{x} - x_j^v, \tag{15}$$

with $x_{j+1}^v = x_i^v + m_i^v =$ for $i = j, ..., J - 1$ and $m_J^v = \overline{x} - x_J^v$.

The interpretation of this result is that under a particular class of adoptions, $v \in \{L, S\}$, the solution can be found by simply applying the mapping $\Gamma^v(., .) = 0$ recursively as in Eq. (14). The restricted solution to the problem in (7) is given when $j = 1$ and $x_1^v = 0$. Here $e^{-rm}\Gamma^v(m, m')$ is shorthand notation for the first-order condition in Eq. (10). Uniqueness and existence are due to the fact that the objective is well behaved and guarantees that $\Gamma^v(., .)$ is monotonic in the choice variables at each stage of the recursion in Eq. (14), and that the value functions are differentiable. The proof uses induction on the fact that these properties hold for $j = J - 1$. Observe that the restricted S-plan is, in fact, the unrestricted optimal plan when there is no scope for learning [i.e. $\mu > T$ or $\delta = 1$.].

Proposition 1 provides an algorithm for solving the optimal v-plan restricted to the number of adoptions being $J - (j - 1)$. Appendix B describes the practical procedure to perform this computation.
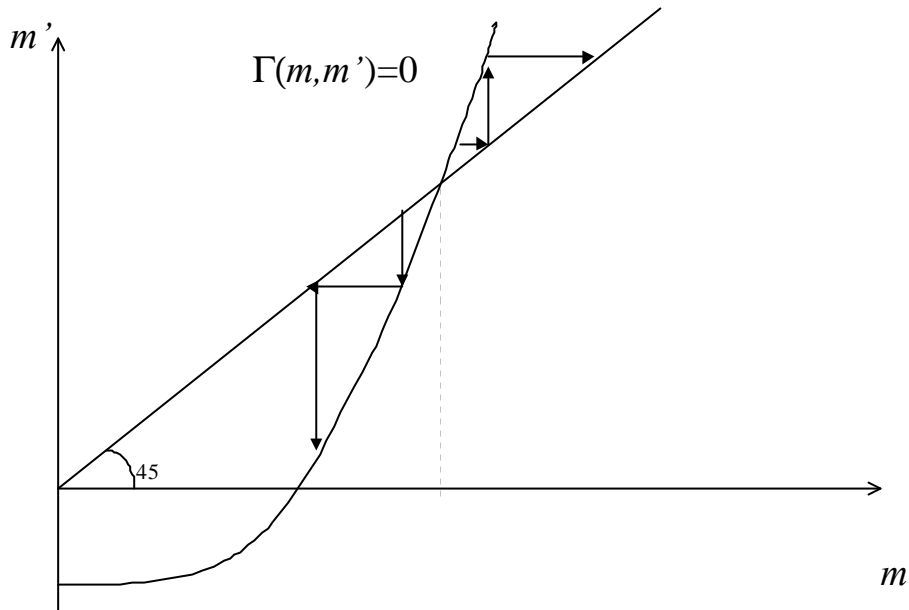
Figure 1. The Dynamics of Adoptions

Figure 1: The curve represents the optimal relation between the tenures in two consecutive adoptions.

The properties of the mapping defined by $\Gamma^v(.,.) = 0$ in Eq. (14) can be analyzed. It is possible to show that the mapping $m' \to m$ is increasing and, as long as $\pi > 0$, has one fixed point which is unstable. Figure 1 shows the typical shape for this mapping. It is clear that the time pattern for tenures depends on the value of the tenure on the first technology relative to this fixed point. But the value of the initial tenure has to be consistent with the constraint in Eq. (15) being satisfied after exactly J steps. In addition, the value of J that is optimal is still to be determined. This is something a priory theorizing cannot resolve.

A reference benchmark is the case where the time horizon for the problem is a multiple of the tenure length that characterizes the fixed point of this mapping. The proof of proposition 1 shows that the solution to Eq. (14) is unique [i.e. $m(.|k)$ is a decreasing function] so, in this case, the fixed point is a solution to the restricted problem if the time horizon contains this span of time exactly J times. From this benchmark, a reduction in J or an increase in T will then tend to increase the initial tenure length and produce an uneven pattern of increasing tenures over time, with lower frequency of adoptions as time goes by. So the pattern of tenures will depend on the features of the optimal choice of

J relative to the one that leads close to the constant profile for tenures. The following proposition characterizes the optimal number of adoptions.

> **Proposition 2.** *Suppose there exists an optimal v-plan over the interval $[\underline{x}, \overline{x}]$, then:*
>
> (i) *There exists a unique sequence $\{z_j^v\}_{j=-\infty}^{J}$ defined by*
>
> $$V^v(z_j^v, \overline{x}|J-j+1) = V^v(z_j^v, \overline{x}|J-(j+1)+1), \qquad (16)$$
>
> *for $j = J-1, J-2, ...$ and with z_J^v such that $W^v(z_J^v, \overline{x}) = 0$.*
>
> (ii) *$\underline{x} = x_{J-k}^v$ if and only if $\underline{x} \in (z_{J-(k+1)}^v, z_{J-k}^v]$ and so $J = k+1$ and*
>
> $$V^v(\underline{x}, \overline{x}) = V^v(\underline{x}, \overline{x}|J).$$

Part (i) determines a sequence of dates where the constrained optimal value of making $k+1$ adoptions is the same as that of k adoptions [in this case $k = J - j$]]. The idea is as follows. There is an early date such that the lifespan is long enough that making a large number of adoptions such as $k+1$ implies a higher present value than making a smaller number of adoptions such as k. However the value of making $k+1$ adoptions relative to the one from k adoptions declines as time draws on and the time horizon becomes shorter. One can show that there is a point in time when the two values are the same, and making one less adoption produces a higher value afterwards. Therefore, such a point z_{J-k}^v constitutes an upper bound for the dates where making $k+1$ adoptions can possibly be optimal. This is illustrated in figure 2 below. Part (ii) of the proposition shows that these points z_j^v are indeed the ones that define the partition on the real line that can be mapped into the optimal number of adoptions.

The procedure to solve for the optimal v-plan is thus as follows: (1) compute the sequence of z_j^v's as in Proposition 2-i, (2) locate the starting date and determine the number of adoptions, J, as in Proposition 2ii, and, finally, (3) use Proposition 1 to calculate the timing of adoptions. Note that computing the z_j^v's in the first step one must already use proposition 1 and a convergent algorithm to find the point where the equality of value functions in Eq. (16) holds. I find that a Newton-Rapson procedure works well.
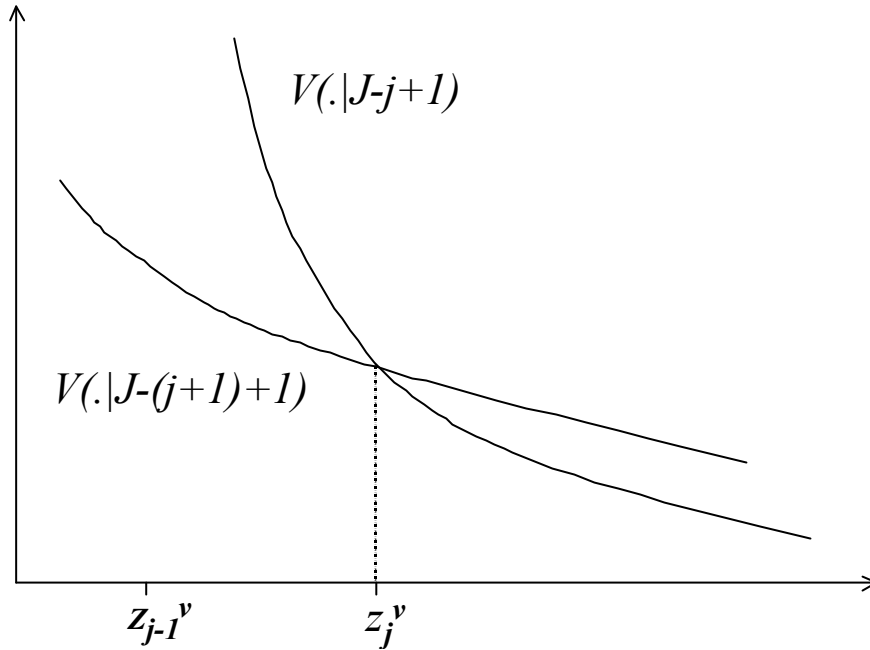
13

Figure 2. The Number of Adoptions

Figure 2: The determination of the number of adoptions.

4.2   Characterization of the optimal plan

This section shows that the previous analysis is useful to compute the solution of the (unrestricted) optimal plan. It is intuitive that the optimal (unconditional) plan over $[0, T]$ contains some L-subinterval if δ is sufficiently large, or μ or γ are sufficiently small, or π is sufficiently large. When circumstances are the opposite, one would expect the optimal plan to contain S-subintervals. There are situations where the entire optimal plan consists of a v-plan for either v. In general, however, the optimal plan may contain L-plans and S-plans over different periods.

Under some circumstances, the optimal plan can be shown to belong to a particular restricted class. A trivial case is that were the learning period exceeds the given lifespan or learning just cannot produce a positive value. Then only a S-plan can be optimal [this are Propositions A1 and A2i in Appendix A]. As long as an adoption with learning can produce a positive value [i.e. z_J^L larger than initial date] then the optimal plan will contain L-adoptions over some interval. If, in addition, a net positive value on a technology requires learning then the optimal plan consists of the restricted L-plan [proposition A2iia].

The case when an adoption without learning can produce a positive value is consistent

14

with the optimal plan containing both L and S plans over different intervals. In these situations, the solution procedure relies on an educated conjecture.

Assumption 3. *The restricted value functions $V^L(x, x+m)$ and $V^S(x, x+m)$ do not intersect more than twice as functions of m.*

This conjecture implies that, actually, the two value functions intersect only once for a second intersection must necessarily imply a third one. The reason is that there is always a span of time long enough that learning the adopted technologies dominates [this is Lemma A4 in Appendix A]. Assumption 3 has to be made explicit because the non-linearities in the restricted value functions preclude to state it as a property. In all the calculations performed in this research this property holds. Under assumption 3, one can argue that, if the optimal plan contains both adoptions of duration longer than μ and adoptions of duration shorter than μ, then the former type of adoptions must occur first. Proposition 3 states this result more precisely and summarizes the discussion thus far.

Proposition 3. *Assume assumption 3 above holds, then the optimal plan solves the following program,*

$$V(0, T) = \max_{x^\star \in [0,T]} \left\{ V^L(0, x^\star) + e^{-(r-\gamma)x^\star} V^S(x^\star, T) \right\}. \tag{17}$$

As a practical concern, searching for the solution without further constraints on the choice set for $x^\star$ is highly inefficient. Propositions A2 to A4 in the appendix identify conditions for which the optimal plan is either the optimal S-plan [i.e. $x^\star = 0$] or the optimal L-plan [i.e. $x^\star = T$], and, otherwise, provide results that narrow down the region where $x^\star$ may lie.

Whereas computation is feasible and efficient, proposition 3 cannot be used to study analytically the pattern of the optimal choice of J and the frequency of technology adoptions. Therefore, these implications will be analyzed numerically.

5   Numerical results

The sequence of steps in propositions A2-A4 in the appendix provide an algorithm to calculate the optimal plans. The algorithm is complete if assumption 3 is verified. In this

section, this procedure is used to illustrate the pattern of technology adoption and assess the role of the finite horizon. We are interested in studying the pattern of tenures over time. There are two reasons why tenures may not be constant. The first, already pointed in the discussion of section 4.1, is that within a restricted plan departures may be expected from the fixed point in figure 1. The second source is the possibility that the solution contains both long adoptions with learning and short adoptions without learning. That is, x^* in the problem of Eq. (15) above may be an interior solution. In this case, more frequent adoptions should be observed towards the end of the period.

The parametric benchmark is $r = 0.065$, $\gamma = 0.02$, $\delta = 2.0$, $\mu = 0.7$, $\pi = 1.12$, and $T = 60$. The figures for r and γ are consistent with observations for the annual real rate of return on equity and aggregate economic growth over long periods. A progress ratio δ is a choice made in other studies on learning-by-doing like Klenow (1998). The speed of learning μ is as calibrated in Mateos-Planas (forthcoming). The time horizon corresponds to 60 years. For this setting, the optimal adoption plan features 6 adoptions and frequency increases over time as m_j declines with $j = 1, ..., 6$. This plan is the optimal L-plan. I will analyze the effect of the parameters on the adoption plan by considering departures from this benchmark setting.

Consider first the parameters characterizing learning-by-doing μ and δ. A reduction in the speed of learning-by-doing is represented by a higher value for μ. Graphically, such a change brings about a downward shift of the curve in figure 1. Assume first that the optimal J remains unaffected. It should be expected that the fixed point will move to the right relative to the value of the initial tenure, thereby tending to increase the frequency of late adoptions relative to that of early ones. The examples computed are consistent with this. In figure 3 below, for $\mu = 0.35$ the path for tenures is increasing rather than decreasing, so that, relative to the benchmark, adoptions become more frequent for earlier periods as μ is reduced. However, a rise in μ will also tend to reduce the number of adoptions J. In this case, a higher μ can make late adoptions relatively less frequent. For example, figure 3 also shows an increasing path for tenures associated with $\mu = 0.85$ and one less adoption.

A rise in the progress ratio δ also shifts downwards the curve in figure 1. Thus, for given J, the slope of the time profile for tenures decreases and, consequently, early adoptions become less frequent and later adoptions become more frequent. If an increase in δ also increases the number of adoptions the contrary effect can be observed. Figure 4 illustrates
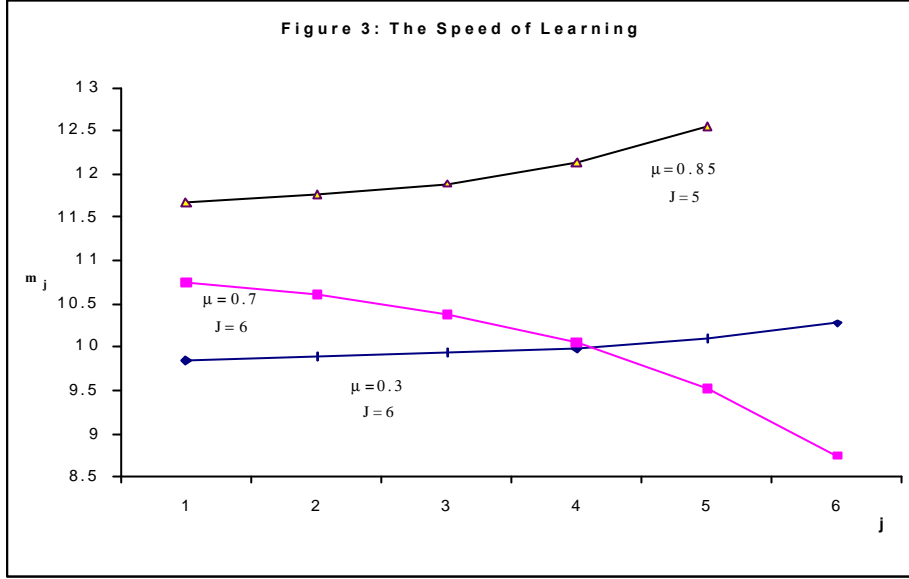
16

Figure 3: The effect of changes in μ on the pattern of adoptions.

the effect of δ on the pattern of optimal adoptions.

For given J, a shorter time horizon T reduces the length of the periods between adoptions and, according to figure 1, tends to reduce the frequency of early adoptions and increase the frequency of late adoptions. When lower T leads to a reduction in the number of adoptions, the effect may be overturned. Figure 5 illustrates this point.

For given J, a higher interest rate r reduces the frequency of early adoptions and increase the frequency of late adoptions. When higher r leads to a reduction in the number of adoptions, the effect may be overturned. Figure 6 illustrates this point.

For given J, a higher rate of technical progress γ has opposite effects on the timing of adoptions: frequency increases in the early adoptions. A sufficient rise in γ makes it optimal to adopt a larger number of technologies. Illustrative paths are shown in figure 7 below.

For given J, the sunk cost of technology adoption, π, reduces the frequency of early adoptions and increases that of late adoptions. A sufficiently large increase in π leads to a smaller number of adoptions. Some illustrative figures are displayed in figure 8.

In the examples reported, the optimal plan is always a L-plan. However, for other parametric settings the optimal plan combines adoptions with and without learning-by-doing [for example, if $\mu = 15.0$]. No example has been found that violates assumption 3 which is reassuring about the general applicability of the solution algorithm developed in
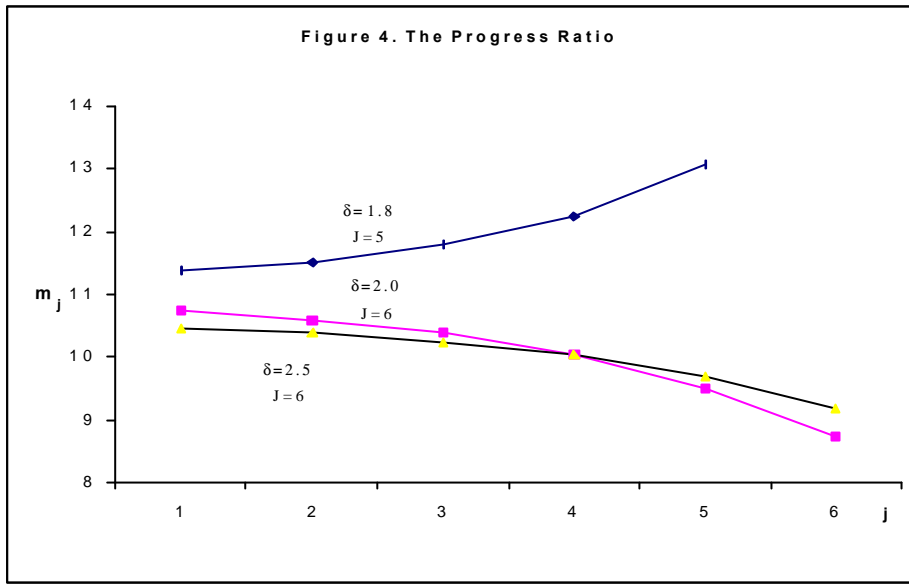
Figure 4: The effect of changes in δ on the pattern of adoptions.
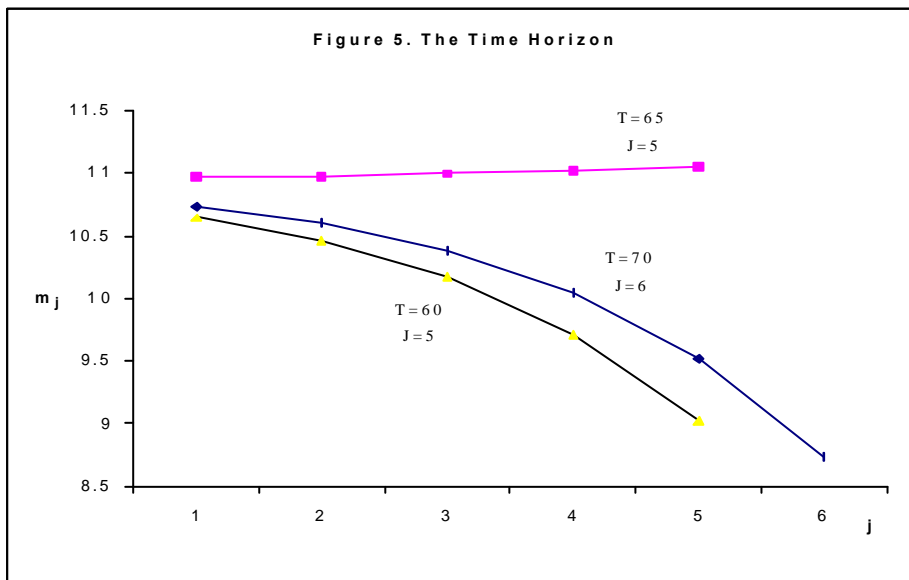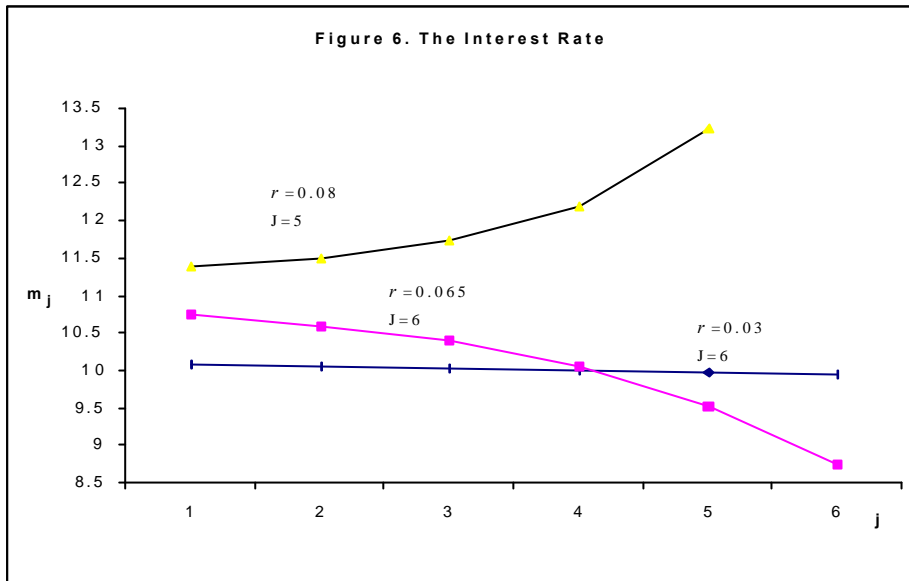

Figure 5: The effect of changes in T on the pattern of adoptions.

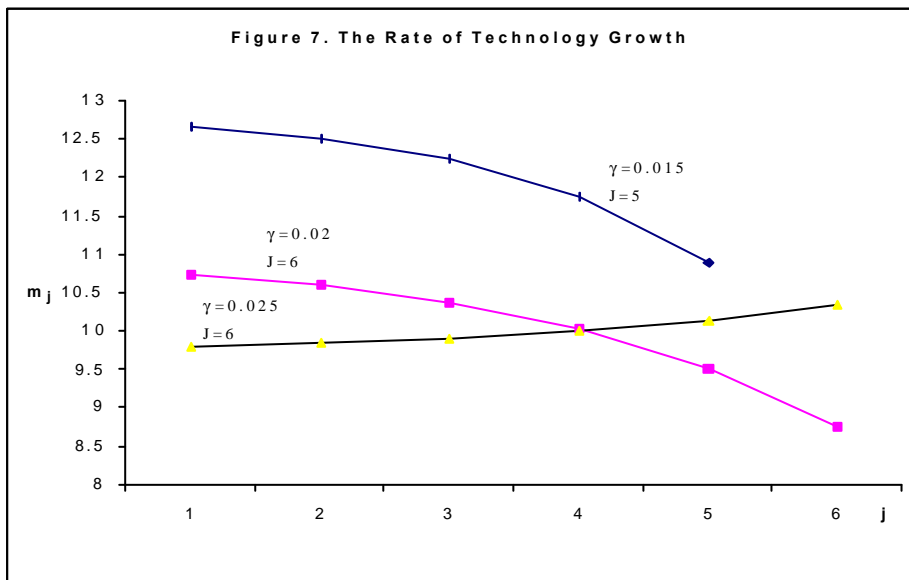Figure 6: The effect of changes in r on the pattern of adoptions.

Figure 7: The effect of changes in γ on the pattern of adoptions.

19

Figure 8. The Adoption Cost

$\pi = 1.5$
$J = 5$
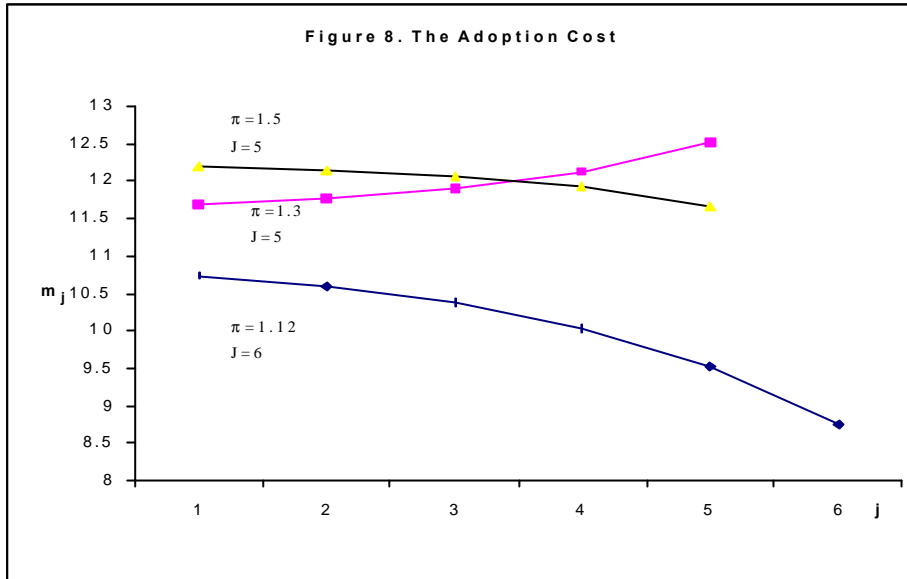
$\pi = 1.3$
$J = 5$

$\pi = 1.12$
$J = 6$

Figure 8: The effect of changes in π on the pattern of adoptions.

this paper. In all cases, like in Klenow (1998), technology upgrading is followed by a drop of productivity.

6   Conclusion and Final Remarks.

This paper analyzes the optimal sequence of technology upgrades by a firm that lives for a finite period of time. Other characteristics of the environment are the existence of technology specific learning-by-doing, technology growth, and sunk costs of technology adoption. The finite planning horizon implies that the problem is non-stationary and the frequency of adoptions changes over time. This paper provides results for the computation of the optimal plan.

The paper also analyzes some properties of the solution. The length of the time period between adoptions is monotonic. The response of the pattern of technology switches to changes in the model's parameters has been explored numerically. The effects of local changes that do not alter the number of switches is clear. Early adoptions become more frequent (and thereby late adoptions become less frequent) the faster the learning-by-doing process, the smaller the adoption cost, the smaller the interest rate, the higher the progress ratio, the longer the time horizon, and the higher the rate of technology growth. However, since the number of switches itself is a choice variable, the model's predictions are in general

20

ambiguous. The assessment of the significance of the finite-horizon approach — relative to approaches based on, for example, on-the-job search as in Topel and Ward (1992)— to interpret actual patterns of technology adoption then must await future work. In a quantitative setting, this work should derive implications that can be compared with the patterns observed in the data.

References

[1] Argotte, L. and Epple, D. Learning curves in manufacturing. *Science, 247*, 1990, 920-924.

[2] Bahk, B. and Gort, M. Decomposing learning by doing in new plants. *Journal of Political Economy, 101*, 1993, 561-583.

[3] Balcer, Y. and Lippman, S.A. Technological Expectations and Adoption of Improved Technology. *Journal of Economic Theory 34*, 1984, 292-318.

[4] Bartel, Ann P. and Frank R. Litchtenberg. The comparative advantage of educated workers in implementing new technology. *The Review of Economics and Statistics, 69, 1*, February 1987, 1-11.

[5] Cooley, Thomas F., Jeremy Greenwood and Mehmet Yorukoglu. The replacement problem. *Journal of Monetary Economics, 40*, 1997, 457-499.

[6] Dixit, A.K. and Pindyck, R.R. *Investment under Uncertainty*, Princeton University Press, 1994.

[7] Doms, Mark, Timothy Dunne and Kenneth R. Troske, Workers, wages, and technology. *Quarterly Journal of Economics*, February 1997, 253-290.

[8] Greenwood, J. and Yorukoglu, M. 1974. *Rochester Center for Economic Research, Working Paper No. 429*, September 1996.

[9] Hendricks, L. Equipment investment and growth in developing countries. *Mimeo*, March 1997.

[10] Jensen, R.A. Adoption and Diffusion of Innovations under Uncertainty. *Journal of Economic Theory 27*, 1982, 182-193.

[11] Jovanovic, B. and Rafael Rob. Solow vs Solow. *Mimeo*, September 1998.

[12] Jovanovic, Boyan. Learning and growth. *NBER Working Paper No. 5383*, December 1995.

[13] Jovanovic, Boyan and Yaw Nyarko. Learning-by-doing and the choice of technology. *Econometrica, Vol. 64, No. 6*, 1996, 1299-1310.

[14] Jovanovic, Boyan and Yaw Nyarko. A Bayesian learning model fitted to a variety of empirical learning curves. *Brooking Papers of Economic Activity (Microeconomics)*, 1995, 245-300.

[15] Kamien, M.I. and Schwartz, N.L.. Timing of Innovations under Rivalry. *Econometrica, 40*, 1972, 46-60.

[16] Klenow, Peter J. Learning curves and the cyclical behavior of manufacturing industries. *Review of Economic Dynamics, 1,2* , April 1998, 531-550.

[17] McDonald, R. and Siegel, D. The Value of Waiting to Invest. *Quarterly Journal of Economics 101*, 1986, 707-728.

[18] Mateos-Plans, Xavier. Schooling and Distortions in a Vintage Capital Model. *Review of Economic Dynamics*, forthcoming.

[19] Parente, Stephen L. Learning-by-Using and the Switch to Better Machines. *Review of Economic Dynamics, 3*, 2000, 675-703.

[20] Parente, Stephen L. Technology adoption, learning-by-doing and economic growth. *Journal of Economic Theory, 63*, 1994, 346-369.

[21] Parente, Stephen L. and Edward C. Prescott. Barriers to technology adoption and development. *Journal of Political Economy, 102, 2*, 1994, 298-321.

[22] Rapping, Leonard A. Learning and World War II production functions. *Review of Economics and Statistics, 47*, 1965, 81-86.

[23] Topel, Robert H. and Michael P. Ward. Job mobility and tye careers of young men. *Quarterly Journal of Economics, May*, 1992, 439-479.

[24] Yorukoglu, Mehmet. The information technology productivity paradox. *Review of Economic Dynamics, 1,2* , April 1998, 551-592.

[25] Zeckhauser, R. Optimality in a world of progress and learning. *Review of Economic Studies*, 1968, 363-365.

A   Proofs of Propositions

The proof of proposition 1 uses the two following lemmas.

Lemma A1. *Consider the adoption plans solving $V^v(x^v_{J-1}, \overline{x}|2)$ for some $v = L, S$.*

(i) *If x^v_{J-1} is the $J-1$th adoption, then x^v_J exists, is unique and satisfies $\Gamma^v(m^v_{J-1}, m^v_J) = 0$.*

(ii) *$m^v_{J-1}(.)$ is a decreasing continuous function.*

(iii) *The value function $V^v(x, \overline{x}|2)$ is continuously differentiable in x with,*

$$\frac{dV^v(x^v_{J-1}, \overline{x}|2)}{dx^v_{J-1}} = W^v_1(x^v_{J-1}, x^v_J) + (r - \gamma)e^{\gamma(x^v_J - x^v_{J-1})}V^v(x^v_J, \overline{x}|1).$$

Proof:

(i) x^v_J is the solution to the problem in Eq. (7) with $j = J - 1$ and $T = \overline{x}$. Clearly, $V^v(x, \overline{x}|1) = W^v(x, \overline{x})$ so the objective is continuous (and differentiable) and the choice set $[x^v_{J-1}, \overline{x}]$ is compact. Therefore a solution exists. A solution must be interior, otherwise x^v_{J-1} cannot be the $J-1$th adoption. The derivative of the objective is $e^{-r(x-x^v_{J-1})}\Gamma^v(x - x^v_{J-1}, \overline{x} - x)$. Since $m^v_J = \overline{x} - x$ is decreasing in x, $\Gamma^v(x - x^v_{J-1}, \overline{x} - x)$ is monotonically decreasing in $x = x^v_{J-1} + m^v_{J-1}$. An interior solution is given by the unique root of $\Gamma^v(x - x^v_{J-1}, \overline{x} - x) = 0$. Clearly it must be a maximum.

(ii) Assume not. As x^v_{J-1} increases, both m^v_{J-1} and, by the properties of $\Gamma(.,.)$, m^v_J must increase. But this violates the constraint. Continuity follows from the continuity of $\Gamma^v(.,.)$.

(iii) Immediate using that the first order condition holds with equality. *Q.E.D.*

Lemma A2. *Consider the adoption plans solving $V^v(x^v_{j+1}, \underline{x}|J - (j + 1) + 1)$ for some $v = L, S$ and some integer $j \quad J - 1$. Assume that,*

(i) *Given x^v_{j+1}, the optimal x^v_{j+2} is unique and satisfies the first order condition, $\Gamma^v(m^v_{j+1}, m^v_{j+2}) = 0$.*

(ii) *$m^v_{j+1}(.)$ is a decreasing continuous function.*

(iii) *The value function $V^v(x, \overline{x}|J - (j + 1) + 1)$ is continuously differentiable and at the optimum,*

$$\frac{dV^v(x^v_{j+1}, \overline{x}|J - (j + 1) + 1)}{dx^v_{j+1}} = W^v_1(x^v_{j+1}, x^v_{j+2}) +$$
$$(r - \gamma)e^{-(r-\gamma)(x^v_{j+2} - x^v_{j+1})}V^v(x^v_{j+2}, \overline{x}|J - (j + 2) + 1).$$

Then the solution of $V^v(x^v_j, \overline{x}|J - j + 1)$ must satisfy the analogous of (i), (ii) and (iii).

Proof:

24

(i) x_{j+1}^v is the solution to the problem in Eq. (7). By assumption (iii), the objective is continuously differentiable and the choice set $[x_j^v, \overline{x}]$ is compact, so a solution exists. A solution must be interior, otherwise x_j^v cannot be the jth adoption.

By assumption (iii), the derivative of the objective is

$$W_2(x_j^v, x_{j+1}^v) + e^{-(r-\gamma)(x_{j+1}^v - x_j^v)} \left[W_1(x_j^v, x_{j+1}^v) + W(x_j^v, x_{j+1}^v) \right],$$

which can also be written as $e^{-r(x-x_j^v)} \Gamma^v(x - x_j^v, m_{j+1}^v(x))$. By assumption (ii), m_{j+1}^v is decreasing and continuous in x, so $\Gamma^v(x - x_j^v, m_{j+1}^v(x))$ is continuous and monotonically decreasing in x. An interior solution is given by the unique root of $\Gamma^v(x - x_j^v, m_{j+1}^v(x)) = 0$ which is a maximum.

ii) By assumption (ii), m_{j+1}^v is decreasing. Inspection of $\Gamma^v(.,.)$ concludes the proof.

iii) Immediate using that the first order condition holds with equality.$Q.E.D.$

Proof of Proposition 1. Lemma A2 says that if properties (i), (ii) and (iii) hold for $V_{J-k}^v(.,.|k+1)$ for some k, then they also hold for $V^v(.,.|k+2)$. Lemma A1 states that these properties hold for $V^v(.,.|k+1)$ for $k=1$. Induction on k then completes the proof by showing that the first order condition in (i) is satisfied for all $k.Q.E.D.$

For the proof of proposition 2 some definitions are required. Consider plans over $[\underline{x}, \overline{x}]$. For L-plans [i.e. $v = L$], the latest possible date for the last adoption, x_J^{LU}, is given by $W^L(x_J^{LU}, \overline{x}) = 0$. For S-plans [i.e. $v = S$], x_J^{SU} is given by $W^S(x_J^{SU}, \overline{x}) = 0$.

Let the sequences $\{m_j^{vU}\}$ and $\{x_j^{vU}\}$ for $j = J, J-1, J-2, ...$ define upper bounds for earlier adoptions j J. These values are found by iterations on $\Gamma^v(.,.)$: $m_J^{vU} = \overline{x} - x_J^U$, $\Gamma(m_{J-1}^{vU}, m_J^{vU}) = 0$, $m_{J-1}^{vU} = x_J^U - x_{J-1}^U$, $\Gamma(m_{J-2}^{vU}, m_{J-1}^{vU}) = 0$, $m_{J-2}^{vU} = x_{J-1}^U - x_{J-2}^U, ...$

Lemma A3. *The expression $\exp\{-(r-\gamma)x\}[V^v(x, \overline{x}|J-(j-1)-1) - V^v(x, \overline{x}|J-j+1)]$ is a continuous decreasing function of x.*

Proof. The proof is divided in two steps. Step 1: Prove that $m_j^v(x) < m_{j+1}^v(x)$. For $j = J-1$ it is obvious. More generally, assume that $m_j^v(x) < m_{j+1}^v(x)$. By Proposition 1, in an optimal v-plan the following must hold:

$$\Gamma(m_{j-1}^v(x), m_j^v(x + m_{j-1}^v(x))) = 0$$
$$\Gamma(m_j^v(x), m_{j+1}^v(x + m_j^v(x))) = 0$$

Suppose, by way of contradiction, that $m_{j-1}^v(x) \geq m_j^v(x)$. Then it must follow that $m_j^v(x + m_{j-1}^v(x)) > m_{j+1}^v(x + m_j^v(x))$. Induction on Lemmas A1 and A2 shows that tenures are non-increasing. This and the assumption made imply $x + m_{j-1}^v(x) < x + m_j^v(x)$. But his contradicts the assumption made. Therefore, $m_{j-1}^v(x) < m_j^v(x)$. Induction on the fact that $m_{J-1}^v(x) < m_J^v(x)$ concludes step 1.

Step 2: The value functions involved in the derivative are continuously differentiable by Lemmas A1 and A.2. Using the expressions obtained there, calculate the derivative of $\exp\{-(r-\gamma)x\}[V^v(x, \overline{x}|J-(j-1)+1) - V^v(x, \overline{x}|J-j+1)]$ as,

$$e^{-(r-\gamma)x} \left(1 - \frac{r-\gamma}{r}\right) hq \left[e^{-rm_j^v(x)} - e^{-rm_{j-1}^v(x)} \right] < 0,$$

with $q = 1$ if $v = S$, and $q = \delta$ if $v = L$. The inequality follows from the result in step 1. Q.E.D.

Proof of proposition 2:

(i) For any $v = L, S$, existence of plans implies existence of x_j^{vU}. The proof contains three steps. Step 1: By definition of x_{J-1}^{vU}, we have that $V^v(x_{J-1}^{vU}, \overline{x} | J - (J-1) + 1) < V^v(x_{J-1}^{vU}, \overline{x} | J - J + 1)$. From Lemma A3, $\exp\{-(r-\gamma)x\}[V^v(x, \overline{x} | J - (J-1) + 1) - V^v(x, \overline{x} | J - J + 1)]$ is monotonically decreasing in x. Therefore, $z_{J-1}^v < x_{J-1}^{vU}$ exists. Step 2: prove that $V^v(x_j^{vU}, \overline{x} | J - j + 1) < V_{j+1}^v(x_j^{vU}, \overline{x} | J - (j+1) + 1)$ all j. Assume that $V^v(x_{j+1}^{vU}, \overline{x} | J - (j+1) + 1) < V^v(x_{j+1}^{vU}, \overline{x} | J - (j+2) + 1)$, then

$$
\begin{aligned}
V^v(x_j^{vU}, \overline{x} | J - j + 1) &= W^v(x_j^{vU}, x_{j+1}^{vU}) + e^{-(r-\gamma)m_j^{vU}} V^v(x_{j+1}^{vU}, \overline{x} | J - (j+1) + 1) \\
&< W^v(x_j^{vU}, x_{j+1}^{vU}) + e^{-(r-\gamma)m_j^{vU}} V^v(x_{j+1}^{vU}, \overline{x} | J - (j+2) + 1) \\
&\quad \max_x \left\{ W^v(x_j^{vU}, x) + e^{-(r-\gamma)(x - x_j^{vU})} V^v(x, \overline{x} | J - (j+2) + 1) \right\} \\
&= V^v(x_j^{vU} | J - (j+1) + 1)
\end{aligned}
$$

Induction on the fact proved in step 1 that the property holds for $j = J - 1$ concludes step 2. Step 3: Because $\exp\{-(r-\gamma)x\}[V^v(x, \overline{x} | J - (j-1) + 1) - V^v(x, \overline{x} | J - j + 1)]$ is monotonically decreasing in x, the sequence $\{z_j^v\}$ exists.

(ii) Step 1: Check that $z_j^v < z_{j+1}^v$. More specifically, $z_j^v - z_{j-1}^v \geq m_{j-1}^v(z_{j-1}^v) > 0$. Suppose not: $z_{j-1}^v + m_{j-1}^v(z_{j-1}^v) > z_j^v$. Then,

$$
\begin{aligned}
&V^v(z_{j-1}^v, \overline{x} | J - (j-1) + 1) \\
&= W^v(z_{j-1}^v, z_{j-1}^v + m_{j-1}^v(z_{j-1}^v)) \\
&\quad + e^{-(r-\gamma)m_{j-1}^v(z_{j-1}^v)} V^v(z_{j-1}^v + m_{j-1}^v(z_{j-1}^v), \overline{x} | J - j + 1) \\
&< W^v(z_{j-1}^v, z_{j-1}^v + m_{j-1}^v(z_{j-1}^v)) \\
&\quad + e^{-(r-\gamma)m_{j-1}^v(z_{j-1}^v)} V_{j+1}^v(z_{j-1}^v + m_{j-1}^v(z_{j-1}^v), \overline{x} | J - (j+1) + 1) \\
&V^v(z_{j-1}^v, \overline{x} | J - j + 1)
\end{aligned}
$$

which contradicts the result in part (i) of this proposition.

Step 2: Clearly, if $\underline{x} = x_{J-k}^v$, $\underline{x} < z_{J-k}^v$ and $\underline{x} > z_{J-(k+1)}^v$. The converse also holds, otherwise $z_j^v > z_{j+1}^v$ for some j. But this possibility has been ruled out in step 1. Q.E.D.

The proof of propoasition 3 requires some intermediate results. In what follows I define a v-subinterval as an interval between two adoptions in the optimal adoption plan where it is optimal an v-plan for $v = L, S$.

Proposition A.1. *Suppose that $\overline{x} - \underline{x} < \mu$. (i) If $z_J^S > \underline{x}$ then $V(\underline{x}, \overline{x}) = V^S(\underline{x}, \overline{x})$. (ii) If $z_J^S < \underline{x}$ then $V(\underline{x}, \overline{x}) = 0$.*

Proof: Clearly, there is no room for an L-subinterval since learning never occurs. If, in addition, the net value from using just a single technology is negative (part (i)), then it is optimal not to make any adoption at all. If the net value can be positive (part (ii)) then the optimal plan consists of the optimal S-plan. Q.E.D.

Proposition A.2. *Suppose that $\overline{x} - \underline{x} > \mu$ then:*

(i) If $z_J^L < \underline{x}$: (ia) If $z_J^S < \underline{x}$ then $V(\underline{x}, \overline{x}) = 0$. (ib) If $z_J^S > \underline{x}$ then $V(\underline{x}, \overline{x}) = V^S(\underline{x}, \overline{x})$.

(ii) If $z_J^L > \underline{x}$: (iia) If $z_J^L < \overline{x} - \mu$ then $V(\underline{x}, \overline{x}) = V^L(\underline{x}, \overline{x})$. (iib) If $z_J^L > \overline{x} - \mu$ then the optimal plan may involve either S-subintervals or L-subintervals or both.

Proof: In case i, any feasible L-plan must yield a negative net value, thus there can not be an L-subinterval in any optimal plan. If, in addition, the net value from using just a single technology is negative (ia), then it is optimal not to make any adoption at all. If the net value can be positive (ib) then the optimal plan consists of the optimal S-plan.

In case ii, there is some L-plan that gives a positive net value. If using a technology has a positive net value only after learning occurs (part (iia)) then the optimal plan is the optimal L-plan. Otherwise (part (iib)) some S-plans exist that give a positive net value. *Q.E.D.*

Proposition A.3 below characterizes the solution in case (iib) of proposition A.2. Some previous results and assumption 3 are needed.

Lemma A4. *Consider the restricted optimal value functions $V^S(\underline{x}, \underline{x} + m)$ and $V^L(\underline{x}, \underline{x} + m)$ for $m > 0$. Then, if the two value functions curves intersect at two different m, then there must be a third intersection.*

Proof. The first intersection is at $m = m_a$ with $m_a \geq \mu$ and $V^S(\underline{x}, \underline{x} + m) > V^L(\underline{x}, \underline{x} + m)$ for some $m > m_a$ [this follows from direct inspection of the . Assume a second intersection occurs at $m = m_b$, with $m_b > m_a$. Clearly, by continuity of the restricted value functions, $V^L(\underline{x}, \underline{x} + m) > V^S(\underline{x}, \underline{x} + m)$ for $m \in (m_a, m_b)$ and $V^L(\underline{x}, \underline{x} + m) < V^S(\underline{x}, \underline{x} + m)$ for some $m > m_b$. I want to show that in this circumstances, a third intersection value $m_c > m_b$ must exist. To do this, proceed in a series of steps. Step 1: For any $m > m_b$, the S-plan includes some adoption x_i^S in $(\underline{x} + m_a, \underline{x} + m_b)$. Suppose that $x_i^S \equiv \max\{x_j^S < \underline{x} + m_b\} < \underline{x} + m_a$, so that output under S-plan is constant over (x_i^S, x_{i+1}^S). Because for $x \in (\underline{x} + m_a, \underline{x} + m_b)$ it holds that $V^L(\underline{x}, x) > V^S(\underline{x}, x)$, output on the L-plan must be larger than output from the S-plan at $\underline{x} + m_a$. But then, until $x_{i+1}^S > \underline{x} + m_b$, the L-plan that sticks with the same technology is better than the optimal S-plan. But this is a contradiction. Step 2: Let $m = m_a + m_b$. Then (1) $V^L(\underline{x}, x_i^S) > V^S(\underline{x}, x_i^S)$ since $x_i^S \in (\underline{x} + m_a, \underline{x} + m_b)$, and (2) since $\underline{x} + m - x_i^S < m_b$ and $\underline{x} + m - x_i^S > m_a$, we have $V^L(x_i^S, \underline{x} + m) > V^S(x_i^S, \underline{x} + m)$. Points (1) and (2) imply that $V^L(\underline{x}, \underline{x} + m) > V^S(\underline{x}, \underline{x} + m)$. Step 3: Continuity of the value functions implies that some $m_c < m$ exists such that the two value functions intersect. *Q.E.D.*

Lemma A5. *Suppose assumption 3 in the main text holds. If the optimal S-plan produces a higher value than the optimal L-plan over a certain span of time, then it must be so for any shorter span of time.*

Proof. Assume $V^S(\underline{x}, x) > V^L(\underline{x}, x)$. If for some $x' < x$, $V^S(\underline{x}, x') < V^L(\underline{x}, x')$, since $V^S(.)$ is well-defined there must be an intersection below x'. By Lemma A4, if for $x > x'$ the inequality is reversed, there must be yet another intersection. But this negates assumption 3. *Q.E.D.*

Lemma A6. *Assume that the optimal can include either S-subintervals or L-subintervals or both. If the optimal plan includes a S-subinterval, then it will occur after an L-subinterval.*

Proof. Suppose not: a S-interval of length m^S occurs before and L-interval of length m^L. It must necessarily be the case that $V^L(\underline{x}+m^S,\overline{x}) > V^S(\underline{x}+m^S,\overline{x})$. But then, by Lemma A5, it must be that $V^L(x_j^S,\overline{x}) > V^S(x_j^S,\overline{x})$ for all x_j^S adoption dates occurring over the S-interval. In particular, this holds for $x_j^S = \underline{x}$. But this leads to the contradiction that a S-interval can not be optimal. Q.E.D.

Proof of Proposition 3: *Consider solving for $V(\underline{x},\overline{x})$. A corollary of Lemma A6 is that*

$$V(\underline{x},\overline{x}) = \max_{x^\star}\left\{V^L(\underline{x},x^\star) + e^{-(r-\gamma)(x^\star-\underline{x})}V^S(x^\star,\overline{x})\right\}, \tag{18}$$

provided that in the situations in Propositions A.1 and A.2, where the optimal plan happens to be either v-plan, the solution $x^\star$ is non-interior. Letting, as in Eq. (5), $\underline{x} = 0$ and $\overline{x} = T$ concludes the proof. Q.E.D.

Proposition A.3. *Assume the conjecture holds. Suppose the conditions in Proposition A.2(iib) hold. Then $z_J^S > \overline{x} - \mu$ and:*

   (i) If $z_{J-1}^L > \overline{x} - \mu$, then $V(\underline{x},\overline{x}) = V^S(\underline{x},\overline{x})$.

   (ii) If $z_{J-1}^S < \overline{x} - \mu$, then there must be an L-subinterval in the optimal plan and $x^\star \in \{[\max\{\overline{x}-\mu,\underline{x}+\mu\}, z_J^S]\cup\{\overline{x}\}\}$, where $x^\star$ solves the program in the proof of Proposition 3.

 (iii) If $z_{J-1}^S > \overline{x} - \mu$ and $z_{J-1}^L < \overline{x} - \mu$, then:

     (iiia) If $V^S(\underline{x},\overline{x}) > V^L(\underline{x},\overline{x})$ then $V(\underline{x},\overline{x}) = V^S(\underline{x},\overline{x})$.

     (iiib) If $V^S(\underline{x},\overline{x}) < V^L(\underline{x},\overline{x})$ then there must be an L-subinterval in the optimal plan and $x^\star \in \{[\max\{x_j^L : V^S(x_j^L,\overline{x}) < V^L(x_j^L,\overline{x})\}, z_J^S] \cup \{\overline{x}\}\}$, where $x^\star$ solves Eq. (18) in the proof of Proposition 3.

Proof: It holds that $W^S(z_J^L,\overline{x}) > W^L(z_J^L,\overline{x}) = 0 = W^S(z_J^S,\overline{x})$, where the inequality results from the assumption that $\overline{x} - z_J^L < \mu$, and the two equalities hold by definition of z_J^v for $v = L, S$.

   (i) In any optimal L-subinterval the last adoption must occur at a distance from the ending date less than $\overline{x} - z_{J-1}^L$. If $z_{J-1}^L > \overline{x} - \mu$, then the distance from the last adoption to the end of the period is less than μ. Thus, no L-subinterval can be optimal for there is a S-plan featuring the same timing that yields a higher value.

  (ii) If $z_{J-1}^S < \overline{x} - \mu$, then $V^S(\overline{x} - \mu,\overline{x}) = V^S(\overline{x} - \mu,\overline{x}|1) = V^L(\overline{x} - \mu,\overline{x})$, and for any $x \in (z_{J-1}^S,\overline{x} - \mu)$ it holds that $V^L(x,\overline{x}) > V^S(x,\overline{x}) = V^S(x,\overline{x}|1)$. Then, by Lemma A5, $V^L(x,\overline{x}) > V^S(x,\overline{x})$ all $x < \overline{x} - \mu$. If $x^\star \neq \overline{x}$, optimality requires that $V^S(x^\star,\overline{x}) > V^L(x^\star,\overline{x})$, thus $x^\star > \overline{x} - \mu$. On the other hand, since $\underline{x} < \overline{x} - \mu$, the optimal plan must include some L-subinterval. Therefore $x^\star > \underline{x} + \mu$.

 (iii) If $z_{J-1}^S > \overline{x} - \mu$ and $z_{J-1}^L < \overline{x} - \mu$, then:

    (iiia) Consider first the case that $V^S(\underline{x},\overline{x}) > V^L(\underline{x},\overline{x})$. Now suppose that there is some L-subinterval in the optimal plan. By Lemma A6, this subinterval is $[\underline{x},x']$ for some $x' \quad \overline{x}$. Optimality requires that $V^S(\underline{x},x') < V^L(\underline{x},x')$. By Lemma A5 we know that $V^S(\underline{x},x) > V^L(\underline{x},x)$ all $x \quad \overline{x}$. This is a contradiction.

(iiib) If $V^S(\underline{x}, \overline{x}) < V^L(\underline{x}, \overline{x})$ there must necessarily be an L-subinterval in the optimal plan, otherwise $V(\underline{x}, \overline{x}) = V^S(\underline{x}, \overline{x})$, a contradiction. By Lemma A6, we have that the L-subinterval must precede the S-subinterval [if the latter exists in the optimal plan]. Hence the objective of the maximization problem in the proposition.

If the optimal plan includes a S-subinterval, then optimality requires that $V^S(x^\star, \overline{x}) > V^L(x^\star, \overline{x})$. Thus $x^\star > \max\{x_j^L : V^S(x_j^L, \overline{x}) < V^L(x_j^L, \overline{x})\}$ If the optimal plan does not include a S-subinterval, then $x^\star = \overline{x}$ and $V(\underline{x}, \overline{x}) = V^L(\underline{x}, \overline{x})$. *Q.E.D.*

B Computing the v-plan for given J

Given the initial date x_{J-k}, and the number of adoptions $k+1$, the algorithm to compute the continuation v-plan is the following. Note I suppress indexes v to save notation.

1. Pick an initial value for x_J.

2. Let $S_J = m_J = \tau + T - x_J$ so that $\partial S_J / \partial x_J = \partial m_J / \partial x_J = -1$.

3. Use the first order condition, $\Gamma^h(m_j, m_{j+1}) = 0$, to compute m_j and

$$\frac{\partial m_j}{\partial x_J} = \frac{\partial m_j}{\partial m_{j+1}} \frac{\partial m_{j+1}}{\partial x_J},$$

$$S_j = S_{j+1} + m_j, \quad \frac{\partial S_j}{\partial x_J} = \frac{\partial S_{j+1}}{\partial x_J} + \frac{\partial m_j}{\partial x_J},$$

for $j = J - 1, J - 2, ..., J - k$.

4. If $S_{J-k} + x_{J-k} - x_J$ is far from zero, start again in step 1 with a new x_J updated according to

$$x_J = x_J - \frac{S_{J-k} + x_{J-k} - x_J}{\frac{\partial S_{J-k}}{\partial x_J} - 1}.$$