Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER Journal of Economic Dynamics & Control 28 (2004) 2129–2154
JOURNAL OF Economic Dynamics & Control
www.elsevier.com/locate/econbase

Technology adoption with finite horizons

Xavier Mateos-Planas*

Economics Division, School of Social Sciences, University of Southampton, Highfield, SO17 1BJ, Southampton, UK

Received 4 June 2002; accepted 10 September 2003

Abstract

The purpose of this paper is to study the relation between the age of an agent and her decisions about adopting new technologies. To this end, I analyze the optimal sequence of technology upgrades by an agent who lives for a finite period of time. Other characteristics of the environment are the existence of technology-specific learning-by-doing, technology growth, and adoption 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. Adoptions may become more frequent as the agent grows older. However, the sign of the association between age and the adoption of new technologies is sensitive to variation in parameters. This is not at odds with the findings of recent empirical studies.

JEL classification: O14; O31

Keywords: Learning-by-doing; Technology adoption

1. Introduction

The focus of this paper is on the relation between an individual's age and her decision to adopt new technologies. This is an interesting issue because the impact of certain factors—demographic structure, retirement policies, pension system, to name but a few—on an economy's productivity is bound to depend on the nature of this relationship. There is a widely held view that older individuals are less willing or inclined to change activities. In fact, the negative association between a worker's age and job turnover is a well-established fact at least since Jovanovic (1979), Mincer

* Tel.: +44-2380 595 669; fax: +44-2380 593 858.
E-mail address: fxmp@soton.ac.uk (X. Mateos-Planas).

doi:10.1016/j.jedc.2003.09.002

and Jovanovic (1981) and Topel (1986). Nonetheless, recent empirical work indicates that this impression may prove misleading when it comes to the adoption of new technologies. Borghans and ter Weel (2002) find on U.K. data that computer use does not depend on the worker's age, Friedberg (2003) shows on U.S. data from the CPS for various years that the age pattern of computer use is fairly flat, and Aaronson and Housinger (1999) cannot provide any evidence of technology-related job displacement affecting more older workers in the U.S. It seems then that, to a first approach, a theory of technology adoption—as opposed to a theory of job mobility—should not imply a strong association between age and the rate of adoption of new technologies. Does standard theory meet this restriction?

The primary objective of this paper is to study the life cycle implications of a model of technology adoption which belongs in a class of models used in the literature. The assumption that the agents making the decisions confront a finite lifetime brings a novel life cycle dimension into the analysis. The secondary objective is then to develop results for characterizing and solving the model under this assumption. These results should be of interest of their own when dealing with finite horizons in similar models.

I analyze the technology switching problem of a single agent with a finite horizon in continuous time. There is a process of technological progress which provides the drive for upgrading. There are costs to switching technologies though. On one hand, there is the 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 processes governing the emergence of new technologies and learning-by-doing are deterministic. Regarding the latter, a step-wise process is assumed in that the productivity gain on a given technology accrues only after it has been used for a certain period of time. If a technology is replaced too soon after adoption no learning will be realized. The key problem consists of the choice by a finitely lived agent of technology adoptions in the presence of exogenous embodied technological change and technology-specific learning-by-doing. This paper analyzes the choice of multiple technology adoptions as a non-stationary dynamic programming problem where both the number of adoptions as well as their timing are decision variables. The solution to this problem permits to investigate the properties of the pattern of technology adoptions and the effects of parameters on it.

The results of this paper are as follows. The analysis provides a procedure for the solution of the problem which exploits the analytical features of the model. Two types of situations can be distinguished. In the first type of situation, which holds across the numerical exercises conducted, learning gains are realized in either all or none of the technologies adopted. The frequency or rate of adoption (or the tenure length on a technology) changes monotonically with the age of the agent. The reason for this monotonicity is the positive association which exists between the time spent using the current technology and the next one. In effect, holding on to the current old technology for longer has a cost in terms of the foregone earnings that could have been reaped by switching instead to a new more productive one. Then the optimal decision balances this higher cost of 'investing' in a better technology with a higher return to this investment. This requires that the next better technology also be held for a longer period.

This monotonic relationship between age and the frequency of adoptions may be, in principle, either increasing or decreasing. To see this, I start with the special situation where all the adoptions are evenly spaced over time so that their frequency is constant. One can now imagine the effect of an exogenous change in parameters which, for the sake of the argument, is small enough that leaves the lifetime number of adoptions unaffected. For example, if adoption costs increase then the agent will optimally decide to postpone the date of the next adoption. With the number of adoptions unchanged, that the current technology is held for longer has to imply more frequent adoptions sometime later in life. But then, by the monotonicity of the relation age–frequency just discussed, the association between age and the frequency of upgrades must turn from flat to positive throughout. The reverse argument goes for the equally plausible possibility of a negative relation.

The sign of this relation between age and the frequency of switches is sensitive to the parameters of the learning-by-doing process, the cost of adoption, the interest rate, the rate of technological progress, and the lifetime length. The paper illustrates through numerical examples the effect of each of these factors and demonstrates that the slope and even the sign of the age–tenure relationship do not change monotonically. The direct effect of any one parameter on the sign of the age–tenure relationship may be overturned by the indirect effect through the optimal adjustment of the lifetime number of adoptions. These results indicate that variation in those parameters across individuals may produce the tenuous association age-adoption rate which the empirical studies cited above seem to imply.

In the second type of situation, an optimal plan may include technologies that are learned along with technologies that are replaced before learning occurs. In such cases, the analysis shows that the adoptions where learning occurs must necessarily take place first. The idea is that the agent will abandon technologies faster late in her life cycle when there is not much time left to reap the productivity gains from learning any single new technology and, as an alternative, it is more profitable to stay close to the technology frontier by switching more frequently. This type of situation does not show in the numerical experiments reported though.

There are models in the literature where the search for a good match quality through switching tends to dictate a positive association between age and tenure, which is consistent with evidence on the life-cycle pattern of turnover for workers. This is a characteristic of the important and large literature following the original Jovanovic (1979), including also Harris and Weiss (1984), where the agent keeps on trying out new jobs until a good enough match is found. The key difference is that here there is instead embodied technological progress which weakens the attachment to any particular technology vintage. The latter paper explores the effect of finite horizon in a matching model similar to Jovanovic's but analyzes the single decision to switch irreversibly from a risky to a safe job. In the present paper, if I were similarly to assume exogenously a single switch, we would also find that an old enough agent would never find it profitable to switch so, in this sense, mobility would decline with age. The present paper demonstrates that this effect of a finite horizon does not carry over to explain the entire lifetime pattern of an endogenous number of adoptions.

Formally, the paper also relates to a body of recent literature that analyzes the replacement/adoption of technologies in dynamic settings, which includes Jovanovic and Nyarko (1996), Parente (1994, 2000), Cooley et al. (1997), Greenwood and Yorukoglu (1997), Karp and Lee (2001), and Mateos-Planas (2001). These papers study infinite-horizon problems and the optimal choices are typically characterized by a constant pattern of technology adoption. In the present paper, the planning horizon is instead finite which implies that the policy function is non-stationary. A simplifying feature in the present paper is that learning-by-doing is completely technology specific. Other papers, like Jovanovic and Nyarko (1996), Parente (1994, 2000), and Karp and Lee (2001) accommodate the partial transferability of knowledge across different technologies. Greenwood and Yorukoglu (1997) 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.

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 numerically 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 and operates a single firm over her productive life. The firm produces output using one machine.[1] The flow of output to 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 $a, q \in R_+$. At any instant of time, a firm may either switch to a more advanced technology or continue to use the present one. I call the *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) and Jovanovic and Nyarko (1996).

[1] Hence I will not be able to deal with the effects of impending technology adoption on the capital accumulation decisions of the firm found in Khan and Ravikumar (2002).

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 she operates. To this end, she decides which technology among those available to use at every instant over his productive life, $[0, T]$. The parameters that the agent takes as given are the learning curve $q(\cdot)$, technological progress γ, the interest rate r, the adoption cost π, and the time horizon T. The shape of $q(\cdot)$ 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 $j = 1, \ldots, J, J+1$ representing the dates at which each jth adoption occurs, such that $0 \leqslant x_j < T$, $x_{j+1} > x_j$ for $j = 1, \ldots, J$, and $x_{J+1} = T$.
(iii) A path for the firm's technology $a(t)$ for $t \geqslant 0$ such that $a(t) \leqslant A(t)$, and $a(t)$ is constant for $t \in (x_j, x_{j+1})$ all $j = 1, \ldots, J$.

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

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

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 $V(\underline{x}, \bar{x})$ denote the optimal value of the firm between an initial date $\underline{x}$ and a terminal date $\bar{x}$, divided by the initial level of technology. The maximization problem of the firm is then

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

where

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

and J, $a(\cdot)$, 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 assumptions:

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. These results follow from the assumption that neither the relative adoption costs, π, nor the speed of learning, $q(\cdot)$, 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. 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 without any loss that $a(0) = 1$, this result allows us to rewrite the agent's problem as

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

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, but it also increases the adoption cost and reduces the benefits from learning. As in other papers in the literature, the optimal timing resolves this trade-off. The choice of the number of switches J is a novel feature of the present 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 $j = 1, 2, \ldots, J-1$ 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 $V(x, \bar{x}|k)$ denote the optimal value of a firm that lives between dates x and $\bar{x}$, conditional on the plan containing exactly k adoptions. Thus if k is optimal $V(x, \bar{x}) = V(x, \bar{x}|k)$, where $V(.,.)$ is as defined in Eq. (5) upon letting $x_1 = x$ and $T = \bar{x}$. Then one can write the problem recursively as follows:

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

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 $m(x|k)$ that gives the optimal duration of use, or tenure, of the technology adopted at time x so $x' = x + m(x|k)$.

With these pieces of notation, the problem of the agent in Eq. (5) can be broken down into the following sequence of problems:

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

for $j = 1, \ldots, J-1$. 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 adoption dates $x_{j+1} = x_j + m_j$ for $j = 1, \ldots, J-1$, and $m_J = x_{J+1} - x_J$.

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 requires solving $V(x, T; J-(j-1))$ for different x and $j = 1, \ldots, J$. 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 assumed may be inconsistent with an optimal choice. The second part of the problem is then to find the optimal J as the solution to

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

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 Eqs. (7) and (8) is as follows. For given J, if a solution exists to Eq. (7) it must satisfy

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

for $j = 1, \ldots, J$. 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. The paper will show that in these situations the above Eq. (9) implies that the optimal interior choice of x_{j+1} in problem (7) must satisfy

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

for $j = 1, \ldots, J-1$. This expression has a clear interpretation in terms of the costs and benefit of delaying the date of the next adoption x_{j+1}. This is a second-order difference equation in x_j with initial and terminal conditions $x_1 = 0$ and $x_{J+1} = T$, respectively.

Alternatively, it can also be regarded as a first-order difference equation in the tenure length $m_j = x_{j+1} - x_j$ with $\sum^J m_j = T$. To establish whether this condition is sufficient to characterize a solution, or gives the only solution, the process for technology-specific skill $q(\cdot)$ must be specified.

I assume $q(\cdot)$ 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), and Argotte and Epple (1990). In particular, the following piece-wise learning curve is assumed:

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

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.[2]

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.

Within each class of plans, an arbitrarily given J is not necessarily consistent with optimality in the sense that it may not be possible to find a feasible sequence that satisfies the recursion in Eq. (10). Similarly, for a given initial date there may be several paths governed by the first-order condition in Eq. (10) that are consistent with feasibility and that differ in the number of adoptions J. Numerically finding the number of adoptions by direct application of Eq. (8) may be costly. The result in Proposition 2 allows us to determine exactly the optimal number of adoptions using an alternative approach. 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 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 two types of restricted plans or a combination of the two types of plans. In the latter case, and under what appears in practice to be a fairly mild assumption, the result in Proposition 3 will show that adoptions where learning takes place must occur first. Further results are provided in Proposition 4 that permit to identify conditions where only one class of restricted plan applies throughout or, otherwise, narrow down the region of search for the solution.

[2] Another possible specification for the learning technology is the continuous curve $q(t) = \delta - (1 - \delta)\exp(-\mu t)$. Here δ represents the progress ratio, and μ is a measure of the speed of learning. This learning curve has been used in Parente (1994). I will comment on this case in later footnotes.

4. Optimal adoption of technologies

With the specification of learning-by-doing in Eq. (11), 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 $W(m)$ at $m=\mu$, there may be multiple local extrema at each stage of the recursion in Eq. (7). Therefore, to characterize a solution it proves useful to consider two classes of restricted adoption 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 adoption 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

$$W^v(m)=\begin{cases}\dfrac{1}{r}[1-\mathrm{e}^{-rm}]-\pi & \text{if } v=S,\\[2mm] \dfrac{1}{r}[1+(\delta-1)\mathrm{e}^{-r\mu}-\delta\mathrm{e}^{-rm}]-\pi & \text{if } v=L.\end{cases} \tag{12}$$

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- and S-plans over different subperiods.

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

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

with

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

The next and following propositions are proved in Appendix A.

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

$$\Gamma^v(m_i^v,m_{i+1}^v)=0 \quad \textit{for } i=J-1,\ldots,j \tag{13}$$

and

$$\sum_{i=j}^{J} m_i^v=\bar{x}-x_j^v \tag{14}$$

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

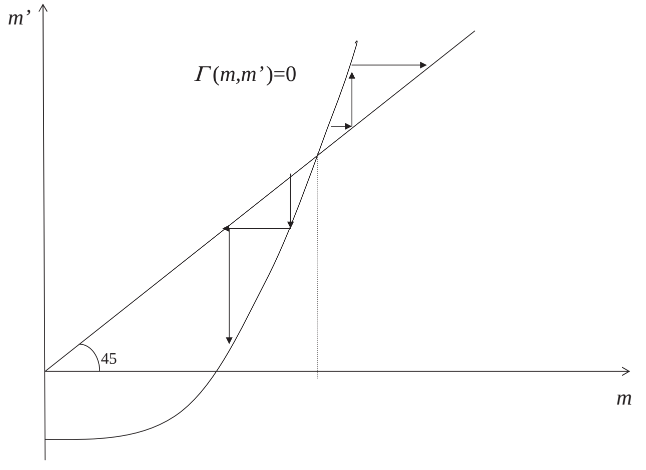

Fig. 1. The dynamics of adoption durations.

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. (13) subject to Eq. (14). The restricted solution to the problem in Eq. (7) is given when $j = 1$, $x_1^v = 0$, and $\bar{x} = T$. Here $\mathrm{e}^{-rm}\Gamma^v(m, m') = 0$ is shorthand notation for the first-order condition to the adoption problem 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. (13), and that the value functions are differentiable. The proof uses induction on the fact that these properties hold for the adoption that precedes the last one, $j = J - 1$. 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.

The properties of the mapping $m' \to m$ defined by $\Gamma^v(.,.) = 0$ in Eq. (13) can be analyzed. It is possible to show that this mapping is increasing and, as long as $\pi > 0$, has one fixed point. Fig. 1 shows the typical shape for this mapping. The positive slope follows from the fact that Eq. (13), being a specific version of the optimality condition Eq. (10), establishes the equality between the marginal gain and loss to increasing the tenure on the current technology, m. The marginal gain in Eq. (10) consists of the flow of output on the current technology which, with the assumed step-wise learning, is constant as represented by either 1 or δ from the term $\Gamma(.,.)$ in Eq. (13). The marginal loss in Eq. (10) consists of two factors. The first is the opportunity cost associated with the discounted foregone present value from a new machine whose quality exceeds the current one's by a factor $\mathrm{e}^{\gamma m}$, which shows explicitly in Eq. (13) through the term $\Gamma(.,.)$. The second factor is the foregone increase in that value which would have accrued if the adoption had taken place earlier rather than later and had thus been used for longer. This factor is represented by $W'(m')$ in Eq. (10) and is more specifically accounted for by the term $e^{-rm'}$ in Eq. (13). The marginal loss then depends positively on the length of the period the current technology is held, m, and negatively on the length of the period the next technology will be held, m', hence the positive sign of the curve's slope. [3]

A monotonic age–tenure relation, either positive or negative, follows directly from the positive relation between any two consecutive tenures just established. The intuition goes along similar lines. Suppose that the second adoption lasts longer than the first. This means that the third adoption is delayed more than the second adoption was and therefore carries a higher investment opportunity cost. Then the third adoption must last longer than the second because, by extending the period over which the gains on this more productive technology can be reaped, the agent matches its higher investment cost. By the reverse argument, if the second adoption lasts shorter than the first, then the third adoption must last shorter than the second.

It is clear from Fig. 1 that the age-pattern of tenures depends on the value of the tenure on the first technology relative to the fixed point depicted in Fig. 1. But the value of the initial tenure has to be consistent with the constraint in Eq. (14) being satisfied after exactly J steps. 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. (13) is unique so in this case the fixed point is a solution to the restricted problem if the planning 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 to a constant tenure length. But the value of J that is optimal is still to be determined. The following proposition characterizes the optimal number of adoptions J.

Proposition 2. *Suppose there exists an optimal v-plan over the interval* $[\underline{x},\bar{x}]$, *then*:

(i) *There exists a unique sequence* $\{z_j^v\}_{j=-\infty}^{J}$ *defined by*

$$V^v(z_j^v,\bar{x}|J-j+1) = V^v(z_j^v,\bar{x}|J-(j+1)+1) \tag{15}$$

for $j = J-1, J-2, \ldots$ *and with* z_J^v *such that* $W^v(z_J^v,\bar{x}) = 0$.

(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*

$$V^v(\underline{x},\bar{x}) = V^v(\underline{x},\bar{x}|J).$$

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, $k = J-j$).

[3] With the smooth learning curve in footnote 1 earlier, the current gain from postponing the adoption of a new technology is no longer constant but increasing, so m raises both the marginal gain and loss. The relation between m and m' can be positive or negative over different ranges and more than one fixed-point could exist.

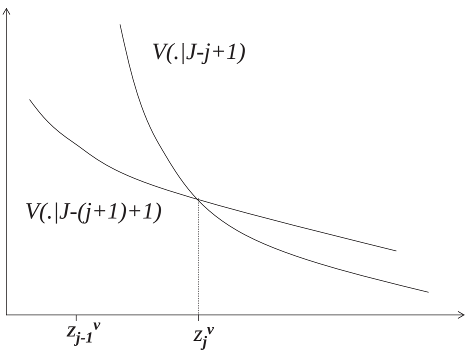

Fig. 2. The determination of the number of adoptions.

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 planning 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^v_{J-k} constitutes an upper bound for the dates where making $k+1$ adoptions can possibly be optimal. This is illustrated in Fig. 2. Part (ii) of the proposition shows that these points z^v_j are indeed the ones that define the partition on the real line that can be mapped into the optimal number of adoptions.

The procedure for solving for the optimal v-plan is thus as follows: (1) compute the sequence of z^v_j's as in Proposition 2(i), (2) locate the starting date and determine the number of adoptions, J, as in Proposition 2(ii), and, finally, (3) use Proposition 1 to calculate the timing of adoptions. Note that computing the z^v_j'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. (15) holds. I have found that a Newton–Rapson procedure works well.

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 plan over $[0, T]$ contains some subinterval where an L-plan is optimal 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 subintervals where a S-plan is optimal. 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. 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 once as functions of m.

Assumption 3 has to be made explicit because the non-linearities in the restricted value functions preclude to state it as a property.[4] 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.

Proposition 3. *If Assumption* 3 *holds, then the optimal plan solves the following program*:

$$V(0,T) = \max_{x^\star \in [0,T]} \{V^L(0, x^\star) + \mathrm{e}^{-(r-\gamma)x^\star} V^S(x^\star, T)\}. \tag{16}$$

Searching for the solution without further constraints on the choice set for $x^\star$ may be highly inefficient. Proposition 4 identifies 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, narrows down the region where $x^\star$ may lie.

Proposition 4. *Let Assumption* 3 *hold. Define the sequences* $\{z_j^v\}$ *as in Eq.* (15) *for* $v = L, S$ *with* $\bar{x} = T$. *Consider the solution* $x^\star$ *to the problem in Eq.* (16). *Assume* $z_J^S > 0$ *so that a solution can exist.*

(a) *If* $T < \mu$ *then* $x^\star = 0$. *If* $T > \mu$ *and* $z_J^L < 0$ *then* $x^\star = 0$. *If* $T > \mu$, $z_J^L > 0$, *and* $z_J^L < T - \mu$ *then* $x^\star = T$.

(b) *If* $T > \mu$, $z_J^L > 0$, *and* $z_J^L > T - \mu$ *then* $z_J^S > T - \mu$ *and*:
- (bi) *If* $z_{J-1}^L > T - \mu$ *then* $x^\star = 0$. *If* $z_{J-1}^S < T - \mu$ *then* $x^\star \in [\max\{T-\mu, \mu\}, z_J^S] \cup \{T\}$.
- (bii) *If* $z_{J-1}^S > T - \mu$ *and* $z_{J-1}^L < T - \mu$ *then*:
  If $V^S(0,T) > V^L(0,T)$ *then* $x^\star = 0$.
  If $V^S(0,T) < V^L(0,T)$ *then* $x^\star \in [\max\{x_j^L : V^S(x_j^L, T) < V^L(x_j^L, T)\}, z_J^S] \cup \{T\}$ *where the* x_j^L*'s belong to the sequence that solves* $V^L(0,T)$.

Hence, under some circumstances the optimal plan can be shown to belong to a particular restricted class. A trivial case is that where the learning period exceeds the given lifespan or learning just cannot produce a positive value. Then only a S-plan can be optimal. As long as an adoption with learning can produce a positive value (i.e., z_J^L larger than the 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

[4] Although it can be proved in the special case that the direct adoption cost is nil, $\pi = 0$. With a positive cost, the trouble is that it becomes hard to compare alternative paths since different timings involve sequences of costs discounted at different rates.

optimal plan consists of the restricted L-plan. In the rest of situations, characterizing the optimal plan requires additional calculations.

5. Numerical results

The sequence of steps in Proposition 4 provides an algorithm for calculating the optimal plan. The algorithm is complete if Assumption 3 is verified. In this section, this procedure is used to study the age pattern of technology adoptions numerically. There are two reasons why tenures may not be constant over the life cycle. The first, already pointed out in the discussion of Section 4.1, is that within a restricted plan departures may be expected from the fixed point in Fig. 1 which may, in principle, go in either direction. 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. (16) above may be an interior solution. In this case, more frequent adoptions should be observed towards the end of the period. The goal of this section is to find out whether in practice these various mechanisms may imply any robust association age–tenure.

The parametric benchmark is $r = 0.065$, $\gamma = 0.02$, $\delta = 2.0$, $\mu = 0.672$, $\pi = 1.12$, 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. The figure for the 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 (2001) consistently with other studies. The adoption cost is also taken from this source where it matches a realistic capital output ratio in a related equilibrium model. The time horizon corresponds to 60 years. For this setting, the optimal adoption plan features six adoptions and their frequency increases over time as tenure m_j declines with $j=1,\ldots,6$. This plan is the optimal L-plan. I analyze the effect of the parameters on the adoption plan by considering departures from this benchmark setting.[5]

In the experiments reported, the sign of the relation age–tenure changes with any one parameter other than the time horizon T through two channels. The first is the impact on the optimal trade-off represented by Eq. (13) which can be visualized by means of shifts in the curve of Fig. 1. More specifically, when this curve shifts upwards, and holding for the moment J fixed, the slope of the relation age–tenure must increase: late tenures become longer relative to early tenures, or late adoptions become less frequent relative to early adoptions. In the current parametric setting, this happens for faster learning or smaller μ, a higher progress ratio δ, a lower discount rate r, a higher embodied technology growth rate γ, or a lower adoption cost π. In the experiments where these changes do not lead to a change in J, this channel can certainly reverse

[5] The continuous learning curve $q(m)=2-(2-1)e^{-1.4566m}$ is 'equivalent' to the discrete process considered here (see Mateos-Planas, 2001) and is also consistent with the speed of learning used in Parente (2000). In spite of the general remarks made in footnote 3, for the benchmark setting this specification preserves the qualitative and quantitative properties of the equilibrium mapping represented in Fig. 1. The crucial point is that learning growth on a single technology dies soon enough. The results of the experiments should be very similar under this smooth curve.

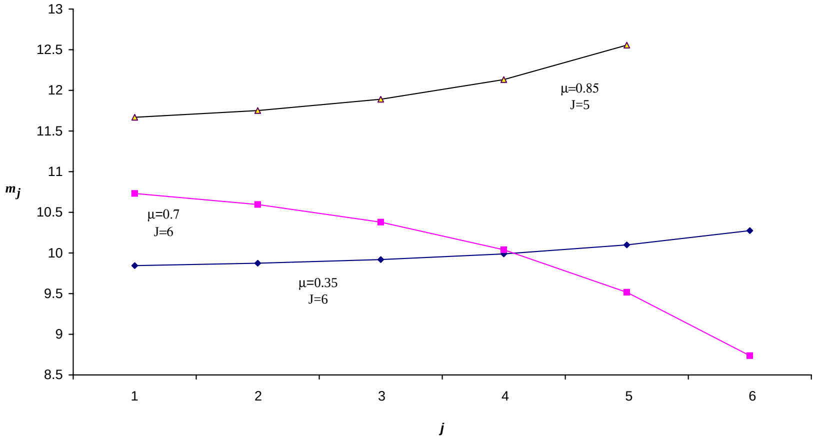

Fig. 3. The time-to-learn μ and the age pattern of adoptions.

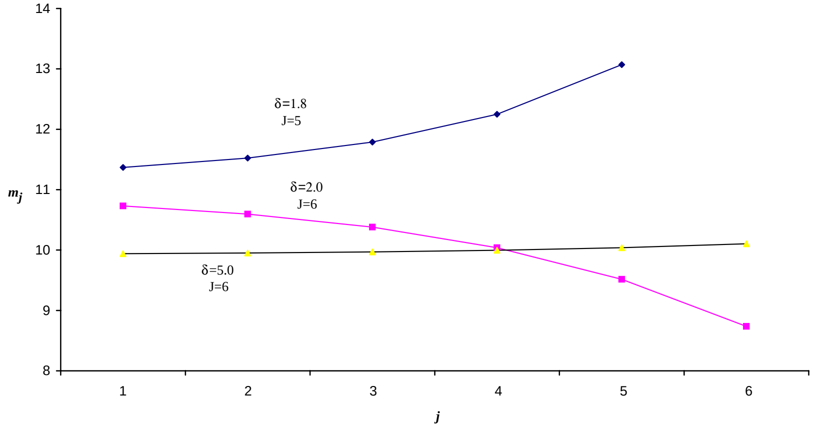

Fig. 4. The progress ratio δ and the age pattern of adoptions.

the benchmark negative association age–tenure. This is demonstrated by comparison of two curves with the same J in Figs. 3–7.

The second channel comes about when parametric changes cause a response of the number of adoptions J. A rise in the number of adoptions will tend to reduce the slope of the age–tenure relationship. This can be seen again in Fig. 1 when time has to be partitioned into more and thus shorter intervals. This effect of increasing J follows in the present setting from a smaller μ, a higher δ, a lower r, a higher γ, or a lower π. Intuitively enough, when learning is faster on a new technology the learning cost of switching is lower, a lower price of the machine embodying the technology also lowers the cost of switching, when new technologies improve more rapidly the return to staying closer to the frontier through switching rises, a lower interest rate increases

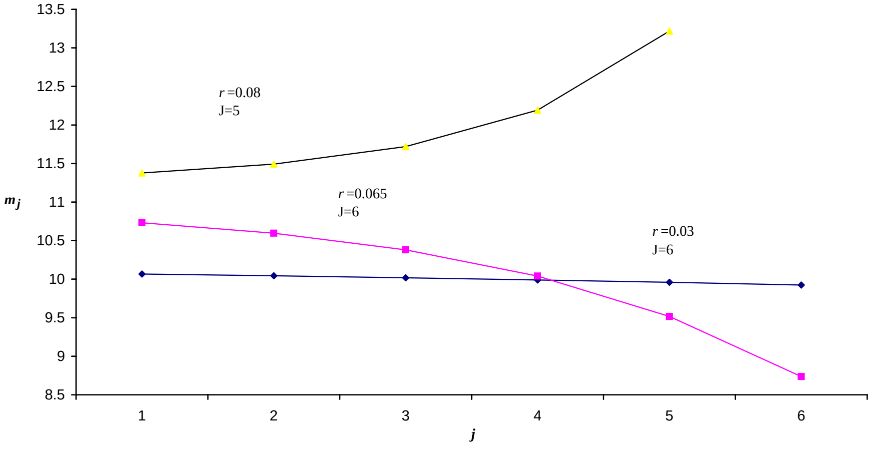

Fig. 5. The interest rate r and the age pattern of adoptions.

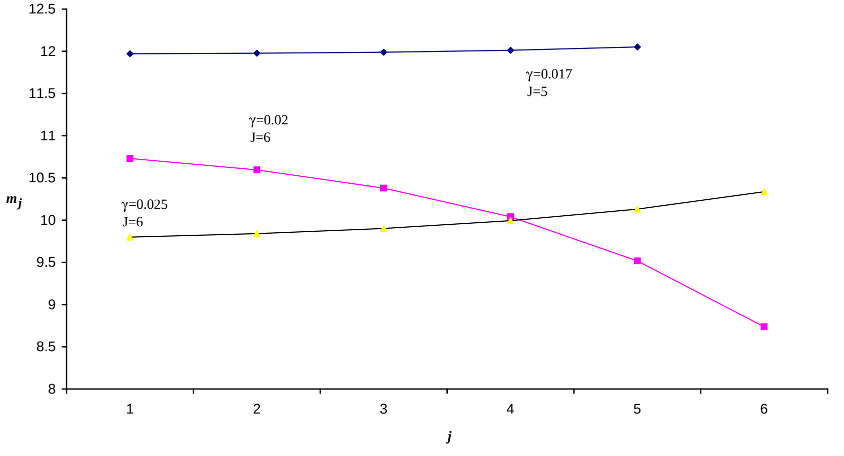

Fig. 6. The rate of technology growth γ and the age pattern of adoptions.

the present value of the future returns from investing in or adopting new technologies. Hence all these factors have a positive effect on the lifetime number of adoptions. The positive effect found of the progress ratio on the number of adoptions is perhaps less intuitive since, while improving the value to adopting and learning a new technology it also makes more costly the loss of specific knowledge. The impact of this channel on the age–tenure relationship in the experiments is demonstrated by comparison or two curves associated with different J's in Figs. 3–7.

An interesting finding follows from this discussion. The two channels just described have opposite consequences on the slope of the age–tenure relationship for each of the parameters μ, δ, r, γ, and π. The experiments reported in Figs. 3–7 demonstrate that

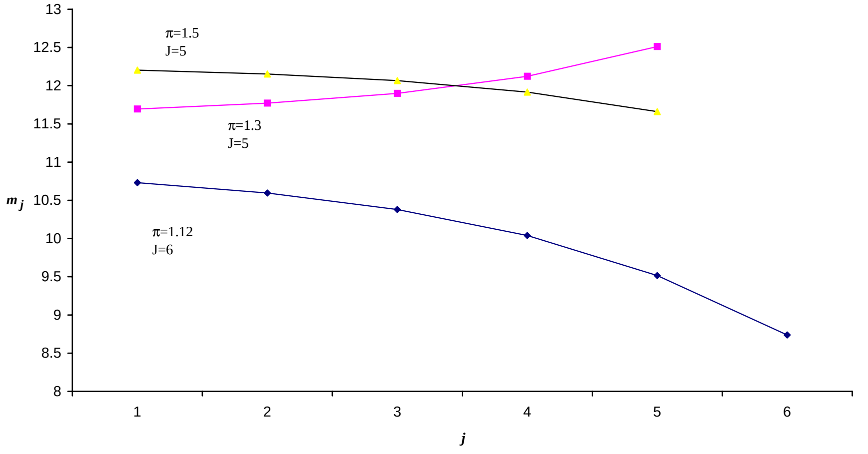

Fig. 7. The adoption cost π and the age pattern of adoptions.

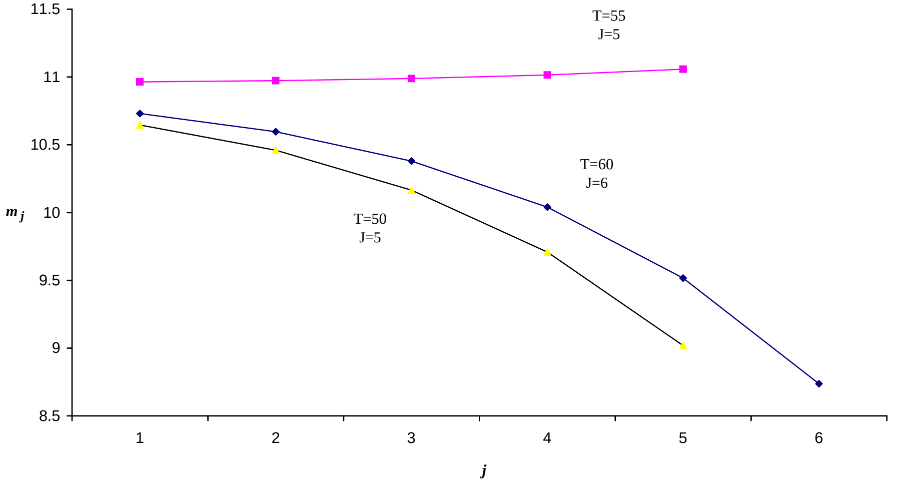

Fig. 8. The time horizon T and the age pattern of adoptions.

variation in any of these parameters may change the sign of this relationship in either direction. It is not only true that the sign of the slope is not robust, but also that the effect of any single parameter on it is largely ambiguous.

Turning now to changes in the lifetime length T, two channels of effects can also be identified. The first is that a longer time horizon, and holding for the moment J fixed, will tend to increase the slope of the age–tenure relationship. This can be seen again in Fig. 1 when time has to be partitioned into longer intervals. This is demonstrated by comparison of the two curves with the same J in Fig. 8. The second channel comes about when a parametric change in T causes a response of the number of adoptions J.

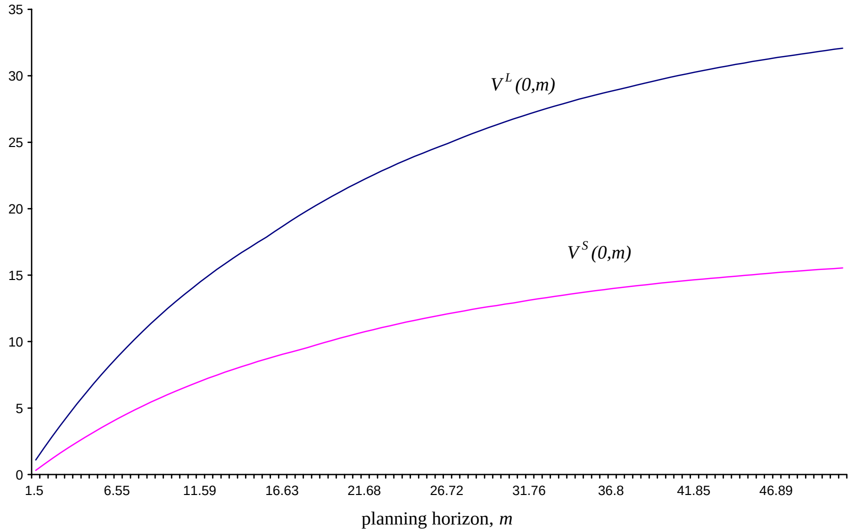

Fig. 9. The value of restricted L and S plans for different horizons.

As already discussed, a rise in the number of adoptions will tend to reduce the slope of the age–tenure relationship, and this will be the result of a sizable enough increase in the time horizon T. Thus, regarding T, the same ambiguous effect found for other parameters also holds. Fig. 8 demonstrates that the sign of the relation age–tenure changes non-monotonically with T.

In the examples reported, the optimal plan is always an L-plan. The possibility of coexistence of two adoption regimes, one where learning occurs and another where it does not, has played no role.[6] Reassuringly, no example has been found that violates Assumption 3. Fig. 9 checks that this assumption holds for the benchmark model, a similar picture holding across the various experiments.

6. Conclusion

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, embodied technology growth, and 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 and analyzes life-cycle properties of the solution.

[6] However, for other parametric settings the optimal plan combines adoptions with and without learning-by-doing. Consider, for example, $\mu = 15.0$ in the otherwise benchmark setup.

The length of the time period between adoptions does generally change over the agent's lifetime monotonically. The response of the age 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 as follows. 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 adoptions is itself a choice variable, the model's predictions for the sign of the age-adoption frequency relation are in general ambiguous. The absence of a strong association between age and technology adoption in the model is not inconsistent with the findings of empirical studies addressing this question.

The evidence on the existence of a link between age and the adoption of new technologies is still fragmentary though. The assessment along these lines of this and other models of technology adoption will have to await firmer empirical findings. In the meantime, the macroeconomic importance that this link might have warrants further theoretical work on the life-cycle of technology adoption.

Acknowledgements

I am grateful to the editor Wouter den Haan and two referees for their thoughtful comments. Early versions have been presented at various conferences and seminars. S.L. Parente made useful remarks.

Appendix A. Proofs of Propositions

The two following lemmas are used in the proof of Proposition 1.

Lemma A.1. *Consider the adoption plans solving* $V^v(x^v_{J-1},\bar{x}|2)$ *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* $\Gamma^v(m^v_{J-1},m^v_J)=0$ *and* $m^v_{J-1}+m^v_J=\bar{x}-x^v_{J-1}$ *with* $x^v_J=x^v_{J-1}+m^v_{J-1}$ *and* $m^v_J=\bar{x}-x^v_J$.
(ii) $m^v_{J-1}(\cdot)$ *is a decreasing continuous function of* x^v_{J-1}.
(iii) *The value function* $V^v(x,\bar{x}|2)$ *is continuously differentiable in* x *with*

$$\frac{\mathrm{d}V^v(x^v_{J-1},\bar{x}|2)}{\mathrm{d}x^v_{J-1}}=W^{v'}(x^v_J-x^v_{J-1})+(r-\gamma)\mathrm{e}^{\gamma(x^v_J-x^v_{J-1})}V^v(x^v_J,\bar{x}|1).$$

Proof.

(i) x^v_J is the solution to the problem in Eq. (7) with $j=J-1$ and $T=\bar{x}$. Clearly, $V^v(x,\bar{x}|1)=W^v(x,\bar{x})$ so the objective is continuous (and differentiable) and the choice set $[x^v_{J-1},\bar{x}]$ is compact. Therefore a solution exists. A solution must be interior, otherwise x^v_{J-1} would not be the $J-1$th adoption. Using the definition of $\Gamma(.,.)$ in Section 4.1, the derivative of the objective can be written as

$e^{-r(x-x^v_{J-1})}\Gamma^v(x-x^v_{J-1},\bar{x}-x)$. Since $\bar{x}-x$ is decreasing in x, $\Gamma^v(x-x^v_{J-1},\bar{x}-x)$ is monotonically decreasing in x. An interior solution is given by the unique root of $\Gamma^v(x-x^v_{J-1},\bar{x}-x)=0$. Clearly, it must be a maximum. The proof is completed upon defining $m^v_i \equiv x^v_{i+1}-x^v_i$ and the feasibility condition that $x^v_{J+1}=\bar{x}$.

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

(iii) This is immediate since the first-order condition holds with equality. □

Lemma A.2. *Consider the adoption plans solving* $V^v(x^v_{j+1},\bar{x}|J-(j+1)+1)$ *for some* $v=L,S$ *and some integer* $j \leqslant J-1$. *Assume that*:

(i) *Given* x^v_{j+1}, *the optimal* x^v_{j+2} *is unique and satisfies the first-order condition,* $\Gamma^v(m^v_{j+1},m^v_{j+2})=0$, *and* $m^v_{j+1}+\cdots+m^v_J=\bar{x}-x^v_{j+1}$ *with* $x^v_{i+1}=x^v_i+m^v_i$ *for* $i=j+1,\dots,J-1$ *and* $m^v_J=\bar{x}-x^v_J$.

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

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

$$\frac{\mathrm{d}V^v(x^v_{j+1},\bar{x}|J-(j+1)+1)}{\mathrm{d}x^v_{j+1}} = W^{v'}(x^v_{j+2}-x^v_{j+1}) + (r-\gamma)e^{-(r-\gamma)(x^v_{j+2}-x^v_{j+1})} \times V^v(x^v_{j+2},\bar{x}|J-(j+2)+1).$$

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

Proof.

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

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

$$W^{v'}(x^v_{j+1}-x^v_j)+e^{-(r-\gamma)(x^v_{j+1}-x^v_j)}[-W^{v'}(x^v_{j+1}-x^v_j)+W^v(x^v_{j+1}-x^v_j)]$$

which can also be written as $e^{-r(x-x^v_j)}\Gamma^v(x-x^v_j,m^v_{j+1}(x))$ using the definition of $\Gamma(.,.)$ in Section 4.1. By assumption (ii), m^v_{j+1} is decreasing and continuous in x, so $\Gamma^v(x-x^v_j,m^v_{j+1}(x))$ is continuous and monotonically decreasing in x. An interior solution is given by the unique root x to $\Gamma^v(x-x^v_j,m^v_{j+1}(x))=0$, which is a maximum. The rest follows from feasibility of the plan and the definition $m_j \equiv x^v_{j+1}-x^v_j$.

(ii) By assumption (ii), $m^v_{j+1}(\cdot)$ is decreasing. Inspection of $\Gamma^v(.,.)$ and feasibility imply that $m^v_j(\cdot)$ must be decreasing as well.

(iii) This is immediate since the first-order condition holds with equality. □

Proof of Proposition 1. Lemma A.2 says that if properties (i), (ii) and (iii) hold for $V^v(.,.|k+1)$ for some k, then they also hold for $V^v(.,.|k+2)$. Lemma A.1 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 so Eq. (13) must hold at each stage of Eq. (7). Feasibility implies Eq. (14). □

The following lemma is instrumental to the proof of Proposition 2.

Lemma A.3. *The expression* $\exp\{-(r-\gamma)x\}[V^v(x,\bar{x}|J-(j-1)-1)-V^v(x,\bar{x}|J-j+1)]$ *is a continuous strictly 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 $m_{j-1}^v(x) \geqslant m_j^v(x)$. Then from the two equations above it follows that $m_j^v(x + m_{j-1}^v(x)) > m_{j+1}^v(x + m_j^v(x))$. Induction on Lemmas A.1 and A.2 shows that tenures $M_j^v(\cdot)$s are non-increasing. This and the assumption made that $m_j^v(x) < m_{j+1}^v(x)$ imply $x + m_{j-1}^v(x) < x + m_j^v(x)$. But this contradicts that $m_{j-1}^v(x) \geqslant m_j^v(x)$. 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 A.1 and A.2. Using the expressions obtained there, calculate the derivative of $\exp\{-(r-\gamma)x\}[V^v(x,\bar{x}|J-(j-1)+1) - V^v(x,\bar{x}|J-j+1)]$ as

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

with $q=1$ if $v=S$, and $q=\delta$ if $v=L$. The inequality follows from the result in Step 1. □

Proof of Proposition 2. Some definitions are needed. Consider plans over $[\underline{x},\bar{x}]$. For a v-plan, the latest possible date for the last adoption, x_J^{vU}, is given by $W^L(x_J^{vU},\bar{x})=0$. Let the sequence $\{x_j^{vU}\}$ for $j=J,J-1,J-2,\ldots$ define upper bounds for earlier adoptions $j \leqslant J$. These values are found by iterations on $\Gamma^v(x_j^{vU} - x_{j-1}^{vU}, x_{j+1}^{vU} - x_j^{vU}) = 0$ for $j=J,J-1,J-2,\ldots$ and $x_{J+1}^{vU} = \bar{x}$.

(i) For any $v=L,S$, existence of an optimal plan implies existence of $x_J^{vU} \in [\underline{x},\bar{x}]$ and thus of z_j^v. The proof contains three steps.

Step 1: By definition of x_{J-1}^{vU}, we have that $V^v(x_{J-1}^{vU},\bar{x}|J-(J-1)+1) < V^v(x_{J-1}^{vU},\bar{x}|J-J+1)$. From Lemma A.3, $\exp\{-(r-\gamma)x\}[V^v(x,\bar{x}|J-(J-1)+1)-V^v(x,\bar{x}|J-J+1)]$ is monotonically decreasing in x. Therefore, $z_{J-1}^v < x_{J-1}^{vU}$ defined in Eq. (15) exists.

Step 2: Prove that $V^v(x_j^{vU},\bar{x}|J-j+1) < V^v(x_j^{vU},\bar{x}|J-(j+1)+1)$ all j. Assume that $V^v(x_{j+1}^{vU},\bar{x}|J-(j+1)+1) < V^v(x_{j+1}^{vU},\bar{x}|J-(j+2)+1)$, then

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

Induction on the fact proved in Step 1 that the property holds for $j=J-1$ concludes Step 2.

Step 3: Because $\exp\{-(r-\gamma)x\}[V^v(x,\bar{x}|J-(j-1)+1)-V^v(x,\bar{x}|J-j+1)]$ is monotonically decreasing in x, the sequence $\{z_j^v\}$ exists for $j=J-2,J-3,\ldots$.

(ii) It follows two steps.

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

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

where the strict inequality follows from the definition of z_j^v and Lemma A.3, which imply $V(x,\bar{x}|J-j+1) < V(x,\bar{x}|J-(j+1)+1)$ for $x > z_j^v$. This contradicts the result Eq. (15) proved in part (i) of this proposition.

Step 2: Clearly, if $\underline{x}=x_{J-k}^v$ then $\underline{x} < z_{J-k}^v$ and $\underline{x} > z_{J-(k+1)}^v$ (see Fig. 2 in the main text for intuition). The converse also holds because otherwise $z_j^v > z_{j+1}^v$ for some j but this possibility has been ruled out in Step 1. □

The proofs of Propositions 3 and 4 are based on the results in Propositions A.1–A.3 and Lemmas A.4 and A.5. It will be useful to define the following.

Definition. A v-subinterval is an interval between any two adoptions in the optimal plan such that a v-plan is optimal for $v\in\{L,S\}$.

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

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. □

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

(i) *If* $z_J^L < \underline{x}$: (ia) *If* $z_J^S < \underline{x}$ *then* $V(\underline{x},\bar{x}) = 0$; (ib) *If* $z_J^S > \underline{x}$ *then* $V(\underline{x},\bar{x}) = V^S(\underline{x},\bar{x})$.
(ii) *If* $z_J^L > \underline{x}$: (iia) *If* $z_J^L < \bar{x} - \mu$ *then* $V(\underline{x},\bar{x}) = V^L(\underline{x},\bar{x})$; (iib) *If* $z_J^L > \bar{x} - \mu$ *then the optimal plan may involve either* S*-subintervals or* L*-subintervals or both.*

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

In part (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-plan exists that gives a positive net value. □

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

Lemma A.4. *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*:

$$V^S(\underline{x}, x) > V^L(\underline{x}, x) \Rightarrow V^S(\underline{x}, x') > V^L(\underline{x}, x') \quad \text{for all } x' < x.$$

Proof. *Step* 1: $V^S(\underline{x},\underline{x}+m) > V^L(\underline{x},\underline{x}+m)$ for all $m < \mu$. This follows because in this case the S-plan gives a higher value than the L-plan for any timing of adoptions.

Step 2: Assume now that $V^S(\underline{x}, x) > V^L(\underline{x}, x)$. If for some $x' < x$ it holds that $V^S(\underline{x}, x') < V^L(\underline{x}, x')$ then the result in Step 1 implies that there must be an intersection below x'. But this negates Assumption 3. □

Lemma A.5. *Suppose Assumption* 3 *holds. Suppose the optimal plan can include both* S*- and* L*-subintervals. The* L*-subinterval precedes the* S*-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 $V^L(\underline{x} + m^S,\bar{x}) > V^S(\underline{x} + m^S,\bar{x})$. But then, by Lemma A.4, it must be that $V^L(x_j^S,\bar{x}) > V^S(x_j^S,\bar{x})$ 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 cannot be optimal. □

Proof of Proposition 3. Consider solving for $V(\underline{x},\bar{x})$. A corollary of Lemma A.5 is that

$$V(\underline{x},\bar{x}) = \max_{x^\star}\{V^L(\underline{x},x^\star) + \mathrm{e}^{-(r-\gamma)(x^\star-\underline{x})}V^S(x^\star,\bar{x})\}, \tag{17}$$

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), $\underline{x} = x_1$ and $\bar{x} = T$, with $x_1 = 0$, concludes the proof. □

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

(i) *If* $z_{J-1}^L > \bar{x} - \mu$ *then* $V(\underline{x},\bar{x}) = V^S(\underline{x},\bar{x})$.
(ii) *If* $z_{J-1}^S < \bar{x} - \mu$ *then there must be an L-subinterval in the optimal plan and* $x^\star \in \{[\max\{\bar{x}-\mu,\underline{x}+\mu\}, z_J^S] \cup \{\bar{x}\}\}$, *where* $x^\star$ *solves the program Eq.* (17) *in the proof of Proposition* 3.
(iii) *If* $z_{J-1}^S > \bar{x} - \mu$ *and* $z_{J-1}^L < \bar{x} - \mu$ *then*:
 (iiia) *If* $V^S(\underline{x},\bar{x}) > V^L(\underline{x},\bar{x})$ *then* $V(\underline{x},\bar{x}) = V^S(\underline{x},\bar{x})$.
 (iiib) *If* $V^S(\underline{x},\bar{x}) < V^L(\underline{x},\bar{x})$ *then there must be an L-subinterval in the optimal plan and* $x^\star \in \{[\max\{x_j^L : V^S(x_j^L,\bar{x}) < V^L(x_j^L,\bar{x})\}; z_J^S] \cup \{\bar{x}\}\}$, *where* $x^\star$ *solves Eq.* (17) *in the proof of Proposition* 3, *and* $\{x_j^L\}$ *is the sequence that solves* $V^L(\underline{x},\bar{x})$.

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

(i) In any optimal L-subinterval the last adoption must occur at a distance from the ending date less than $\bar{x} - z_{J-1}^L$. If $z_{J-1}^L > \bar{x} - \mu$ 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 an S-plan featuring the same timing that yields a higher value.
(ii) If $z_{J-1}^S < \bar{x} - \mu$ then $V^S(\bar{x}-\mu,\bar{x}) = V^S(\bar{x}-\mu,\bar{x}|1) = V^L(\bar{x}-\mu,\bar{x})$, and for any $x \in (z_{J-1}^S, \bar{x}-\mu)$ it holds that $V^L(x,\bar{x}) > V^S(x,\bar{x}) = V^S(x,\bar{x}|1)$. Then, by Lemma A.4, $V^L(x,\bar{x}) > V^S(x,\bar{x})$ all $x < \bar{x} - \mu$. If $x^\star \neq \bar{x}$ optimality requires that $V^S(x^\star,\bar{x}) > V^L(x^\star,\bar{x})$, thus $x^\star > \bar{x} - \mu$. On the other hand, since $\underline{x} < \bar{x} - \mu$, the optimal plan must include some L-subinterval. Therefore, $x^\star > \underline{x} + \mu$.
(iii) If $z_{J-1}^S > \bar{x} - \mu$ and $z_{J-1}^L < \bar{x} - \mu$ then:
 (iiia) Consider first the case that $V^S(\underline{x},\bar{x}) > V^L(\underline{x},\bar{x})$. Now suppose that there is some L-subinterval in the optimal plan. By Lemma A.5, this subinterval is $[\underline{x}, x']$ for some $x' \leqslant \bar{x}$. Optimality requires that $V^S(\underline{x},x') < V^L(\underline{x},x')$. But by Lemma A.4 we know that $V^S(\underline{x},x) > V^L(\underline{x},x)$ all $x \leqslant \bar{x}$. This is a contradiction.
 (iiib) If $V^S(\underline{x},\bar{x}) < V^L(\underline{x},\bar{x})$ there must necessarily be an L-subinterval in the optimal plan, otherwise $V(\underline{x},\bar{x}) = V^S(\underline{x},\bar{x})$, a contradiction. By Lemma A.5 we have that the L-subinterval must precede the S-subinterval.

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

Proof of Proposition 4. This is a corollary of Propositions A.1–A.3 upon letting $\underline{x} = 0$ and $\bar{x} = T$. □

Appendix 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 $S_J = m_J = T - x_J$ so that $\partial S_J/\partial x_J = \partial m_J/\partial x_J = -1$.
3. Use the first-order condition, $\Gamma(m_j, m_{j+1}) = 0$, to compute m_j and
$$\frac{\partial m_j}{\partial x_J} = \frac{\partial m_j}{\partial m_{j+1}} \frac{\partial m_{j+1}}{\partial x_J},$$
$$S_j = S_{j+1} + m_j, \qquad \frac{\partial S_j}{\partial x_J} = \frac{\partial S_{j+1}}{\partial x_J} + \frac{\partial m_j}{\partial x_J}$$
for $j = J-1, J-2, \ldots, J-k$.
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
$$x_J = x_J - \frac{S_{J-k} + x_{J-k} - x_J}{(\partial S_{J-k}/\partial x_J) - 1}.$$

References

Aaronson, D., Housinger, K., 1999. The impact of technology on displacement and reemployment. Economic Perspectives, Federal Reserve Bank of Chicago, second quarter, pp. 14–30.

Argotte, L., Epple, D., 1990. Learning curves in manufacturing. Science 247, 920–924.

Bahk, B., Gort, M., 1993. Decomposing learning by doing in new plants. Journal of Political Economy 101, 561–583.

Borghans, L., ter Weel, B., 2002. Do older workers have more trouble using a computer than younger workers? Research in Labor Economics 21, 139–173.

Cooley, T.F., Greenwood, J., Yorukoglu, M., 1997. The replacement problem. Journal of Monetary Economics 40, 457–499.

Friedberg, L., 2003. The impact of technological change on older workers: evidence from data on computer use. Industrial and Labor Relations Review 56 (3), 511.

Greenwood, J., Yorukoglu, M., 1997. Carnegie-Rochester Conference Series on Public Policy 46, 49–95.

Harris, M., Weiss, Y., 1984. Job matching with finite horizon and risk aversion. Journal of Political Economy 92, 758–779.

Jovanovic, B., 1979. Job matching and the theory of turnover. Journal of Political Economy 87, 972–990.

Jovanovic, B., Nyarko, Y., 1995. A Bayesian learning model fitted to a variety of empirical learning curves. Brooking Papers of Economic Activity (Microeconomics) 1, 245–300.

Jovanovic, B., Nyarko, Y., 1996. Learning-by-doing and the choice of technology. Econometrica 64, 1299–1310.

Karp, L., Lee, I.H., 2001. Learning-by-doing and the choice of technology: the role of patience. Journal of Economic Theory 100, 73–92.

Khan, A., Ravikumar, B., 2002. Costly technology adoption and capital accumulation. Review of Economic Dynamics 5, 489–502.

Klenow, P.J., 1998. Learning curves and the cyclical behavior of manufacturing industries. Review of Economic Dynamics 1, 531–550.

Mateos-Planas, X., 2001. Schooling and distortions in a vintage capital model. Review of Economic Dynamics 4, 127–158.

Mincer, J., Jovanovic, B., 1981. Labor mobility and wages. In: Rosen, S. (Ed.), Studies in Labor Markets. University of Chicago Press, Chicago, pp. 21–63.

Parente, S.L., 1994. Technology adoption, learning-by-doing and economic growth. Journal of Economic Theory 63, 346–369.

Parente, S.L., 2000. Learning-by-using and the switch to better machines. Review of Economic Dynamics 3, 675–703.