Optimal pattern of technology adoption under embodiment with a finite planning horizon: A multi-stage optimal control approach[*]

Hüseyin Çağrı Sağlam[†]

July 22, 2002

Abstract

By deriving the necessary conditions for a multi-stage discounted optimal control problem where the endogenous switching instants between regimes appear as an argument of the objective function and the state equation, we analyze the optimal pattern of technology adoption under embodiment with a finite planning horizon. The economy is characterized by the existence of an exogenously growing technology frontier and technology specific learning by doing. We obtain time varying durations for the adopted technologies to be in use due to finite planning horizon. We analyze numerically the effects of planning horizon, speed of learning, growth rate of technology and impatience rate on the optimal pattern.

Keywords: Multi-stage optimal control, technology adoption, learning by doing, embodiment

Journal of Economic Literature: E22, E32, O40, C63.

---

[*]This paper has benefitted substantially from the comments of Raouf Boucekkine. The support of Belgian research programme "Action de Recherches Concertée" 99/04-235 is gratefully acknowledged.

[†]IRES, Université Catholique de Louvain, Louvain-la-Neuve (Belgium). E-mail: saglam@ires.ucl.ac.be

1 Introduction

The recent economic growth literature has been markedly influenced by the renewed interest in the embodied component of technological progress on the one hand and by the development of adoption-based growth models on the other. Greenwood, Hercowitz and Krussel (1997) have found that around 60% of US productivity growth can be attributed to the embodied technological change. Embodiment is indeed a key concept to understand the economic mechanisms at work since the first oil shock, as argued Greenwood and Yorukoglu (1997). After that, the rate of decline of the relative price of capital has been higher on average and the accumulation of new equipment has been boosted. Both ingredients are proving most useful in understanding and explaining the most interesting and puzzling stylized facts of the real economies, such as the productivity slowdown and the decline in the relative price of capital.[1]

Some important advances have been already accomplished regarding the way technology adoption and the costs associated to it can affect productivity growth and the development process [see the survey of Greenwood and Jovanovic (1998)]. The key aspect of the analysis turns out to be learning: The technology adoption implies the depreciation of the initial human capital and a slow learning process during which the economy is not able to run the new technology at its best productivity level [see Jovanovic (1997) and Parente (1994)]. Indeed, if the technological progress is embodied in the new capital goods and innovations occur continuously, the timing and the nature of new technology adoption rely more heavily on the ability of the economy to learn quickly how to use efficiently the new capital goods.

Despite accumulating knowledge has no direct cost under learning by doing, embodied technological change [as in the case of information technologies] implies an indirect cost due to the obsolescence of the existing capital goods. An acceleration in embodied technical change is associated with a decrease in the relative price of capital [a well known property since Solow (1960)]. This induces a re-assignment of the resources towards capital goods sector resulting in a drop in the consumption level from the date of the technology adoption. The welfare cost due to the drop in the consumption level are referred to as the obsolescence costs inherent to embodiment and they are shown to be non-negligible [see Krussel (1998); Boucekkine, del Rio and Licandro (1999), del Rio (2002)].

In this paper; we analyze the optimal pattern of technology adoption under embodiment when the planning horizon for the economy is finite. Knowing that the heads of governments are elected for a certain period of time, the fixed-term contracts have predetermined terminal date, some stabilization and industrial development programs are assumed for a fixed time interval and the patent protection is guaranteed for a finite period of time, it proves to be important to analyze the optimal pattern of technology adoption with finite horizon as

[1] Benhabib and Rustichini (1991), Greenwood and Yorukoglu (1997) and Krussel (1998) are some examples of these studies in the macroeconomic literature.

well. Other characteristics of the economy are the existence of exogenously growing technology frontier and the technology specific learning by doing which are common in the macroeconomics literature [see Greenwood and Jovanovic (1998), Parente (2000) and Mateos-Planas (2000)]. The level of expertise on a technology evolves with its use as a result of learning-by-doing. Within each generation of technology, productivity increases over time and converges to its potential level with a decelerating growth rate. In case of a switching to a new technology, no part of the current expertise can be transferred for use to the more advanced technology.

In that we consider optimal pattern of technology adoption and learning by doing jointly, our theory resembles works by Jovanovic and Nyarko (1996), Parente (1994, 2000), Iocompetta (2001) and Mateos-Planas (2000). In a partial equilibrium set-up, Jovanovic and Nyarko (1996) assumes a Bayesian learning process so that as long as the economic agents do not estimate correctly some productivity parameters, an output loss occurs. In a general equilibrium framework, Parente (1994, 2000) and Iocompetta (2001) analyze the optimal adoption problem where the learning costs are essentially due to the depreciation of the pre-existing human capital. These papers study infinite horizon problems and conclude that the optimal adoption plans are typically characterized by evenly spaced pattern of technology adoptions. Moreover, in Parente's model, as there is no possibility to stick to a given technology, there is no room to handle technological sclerosis cases. In contrast with these, Mateos-Planas (2000) incorporates a finite planning horizon and predicts a non-stationarity in the sense that an optimal plan may incorporate both types of adopted technologies that are to be learned and that are to be skipped without learning. However, it incorporates a simple discrete learning process and the model's predictions about how the pattern of technology adoption responds to the changes in the parameters are determined to be ambiguous in general.

Our contribution is twofold in this paper. A technical contribution of our analysis is that we extend the analysis of Tomiyama and Rossana (1989) to a multi-stage setting. Tomiyama and Rossana (1989) presents the necessary conditions for a general two-stage optimal control problem with an adjustable single switching time appearance in both the objective function and the state equation. The technique recalls the Pontryagin's maximum principle from a dynamic programming perspective. The approach has been used in the literature first by Tomiyama (1985), for a general two-stage optimal control problem. The analysis has been extended to an infinite-horizon multi-stage optimal control problem by Makris (2001). However, in Tomiyama (1985) and Makris (2001), the problems where the switch-point appears as an argument of the integrands in each integral which compromise the criterion-index to be maximized, have not been considered.[2] In comparison with these, the novel feature of our analysis is to extend the technique to a multi-stage; discounted optimal control problem with finite-horizon, where the switching times being choice variables, appear in

[2] Problems of this variety have also arisen in the exhoustible resource and neoclassical investment literature. See Maccini (1973) and Rossana (1985).

both the objective function and the state equations. The technique proves to be useful and efficient when one would like to analyze the optimal timing of any endogenous regime switches.

Our framework allows us to bring out a number of contributions to the optimal adoption literature. First of all, the optimal pattern of technology adoption will be determined, taking into account not only the eventual learning costs but also the non-negligible obsolescence costs inherent to embodiment. For a given initial state of the economy, including its technological and skill states, we characterize the optimal number of technology upgrades and the optimal timing of them within a given finite planning horizon. In contrast with models by Parente and its extensions, we obtain a non-stationarity in the durations of the adopted technologies to be in use due to the finite planning horizon. The non-stationarity mainly arises at the final technology in use because of the fact that the duration left after the last technology upgrade may not be sufficient to cover the obsolescence and the learning costs of a further upgrade. In contrast with Mateos-Planas (2000), optimal plan does not involve the adoption of a technology which would be skipped without gaining enough experience. The optimal adoption plan involves a number of technology upgrades that ensures within each generation of adopted technology, the elimination of the resulting expertise gap. We illustrate how the optimal pattern of technology adoption evolves with respect to the planning horizon, the speed of learning process, the rate of growth of technology and the discount rate through numerical analysis.

This paper is organized as follows. Section 2 describes the model and the optimal technology adoption problem. Section 3 presents the multi-stage optimal control problem as a resolution process. Section 4 illustrates the optimal pattern of technology adoption and how it responds to the changes in the parameters of the model with a numerical analysis. Finally, section 5 concludes.

2 The Model

We consider an economy inhabited by a representative agent whose discounted stream of utility is given by:

$$\int_0^T u(c(t))\, e^{-\rho t}\, dt; \qquad (1)$$

where u is increasing and strictly concave function, c denotes the flow of consumption and ρ is the time discounting parameter. The planning horizon T, is finite. The economy produces a composite good with a simple AK technology that is used either to consume or to invest in physical capital. Technological progress is investment-specific and it only affects through the new capital goods. Let $Y(t)$; $K(t)$; and $I(t)$ be the output produced, capital and the gross investment respectively at time t. Then, we have the following usual resource constraint and the law of motion of capital stock[3]:

$$Y(t) = A(t)K(t) = c(t) + I(t); \tag{2}$$

$$\dot{K}(t) = q(t)\, I(t). \tag{3}$$

The variable $q(t)$ represents the level of technology in use at time t. An increase in $q(t)$ only affects new equipment by equation (3) and represents embodied technological progress. In sharp contrast, an increase in the variable $A(t)$ which represents the agent's technology-specific expertise, rises the marginal productivity of all the capital stock, independent of its age structure. In this sense, A is neutral and q is investment specific. This simple model is a reduced form of the Solow (1960) vintage capital model as pointed out in Boucekkine, del Rio and Licandro (2000).

At any time, the agent may either switch to a more advanced technology that will lead to a higher rate of investment-specific technological progress or continue to use the current one. We assume that the economy does not innovate and just adopt the technologies coming from abroad. Following Parente (2000), we incorporate a frontier level of technology, denoted by $\bar{q}(t)$. At time t, the technology used by the agent cannot exceed the frontier technology so that $q(t) \leq \bar{q}(t)$. The frontier technology is assumed to grow exogenously at a constant rate $\gamma > 0$, i.e., $\bar{q}(t) = \bar{q}(0)e^{\gamma t}$. Without loss of generality, $\bar{q}(0)$ is normalized to 1.

The level of expertise on a technology evolves with its use as a result of learning-by-doing. Learning in any technology is subject to diminishing returns and is bounded from above by 1.[4] The functional form of learning, which is identical to Parente (1994), is

$$\dot{A}(t) = \mu[1 - A(t)]; \; \mu > 0; \tag{4}$$

where μ is a parameter of speed of learning. Due to this learning effect, the technology adoption incurs a cost in terms of lost expertise. In this model, learning-by-doing is technology-specific and therefore, in case of a switching to a new technology no part of the current expertise can be transferred for use in the more advanced technology.[5] This lost expertise can not be sold or rented. As the amount of lost expertise and the speed of learning do not depend on the productivity of the technology to be adopted, the adopted technology will always be the frontier in an optimal adoption plan. The technology that is currently used influences the time at which the next adoption will be introduced but has no impact on the choice of which technology to adopt at that time. Thus, the agent's choice consists of deciding at any time whether to continue to use the

[3] Note that capital depreciation is omitted for sake of simplicity.

[4] Bounded learning is consistent with the empirical literature as in Bahk and Gort (1993).

[5] Mateos-Planas (2000) also assumes technology specific learning-by-doing. This is a simplification with respect to Jovanovic and Nyarko (1996) and Parente (1994, 2000) without sacrificing the spirit of the model.

current technology or to switch to the frontier technology.[6] If an adoption occurs at time $t = t_s$; then $q(t) = q_f(t_s) = e^{\gamma t_s}$; until the next adoption. Between the two successive adoptions, the technological gap between the technology in use and the frontier one will increase.

2.1 The Technology Adoption Problem

The fundamental decision to be taken by a representative agent in such an economy is as follows: For a given stock of capital K_0; the parameter values of the speed of learning θ; technological progress γ, and the time horizon T; how many numbers of adoptions has to be made and at which dates each of these adoptions should take place for the welfare of the economy to be maximized?

$$\max_{\{c(t);\ J;\ t_1;\ t_2;\ldots;t_J\}} V(T|J) = \int_0^T u(c(t))e^{-\rho t}dt$$

subject to

$$Y(t) = A(t)K(t) \tag{5}$$

$$\dot{K}(t) = \begin{cases} q(t_{j-1})I(t); & \text{if } t_{j-1} \le t < t_j;\ j = 1;\ldots;J \\ q(t_J)I(t); & \text{if } t_J \le t \le T \end{cases} \tag{6}$$

$$A(t) = \begin{cases} 1; & \text{if } 0 \le t < t_1 \\ 1 - e^{-\theta(t - t_{j-1})}; & \text{if } t_{j-1} \le t < t_j;\ j = 2;\ldots;J \\ 1 - e^{-\theta(t - t_J)}; & \text{if } t_J \le t \le T \end{cases} \tag{7}$$

$$Y(t) = c(t) + I(t) \tag{8}$$

The integer j index the j'th technology adopted and the real t_j represents the time of that adoption. Note that the cost of switching limits the number of adoptions (J) within a given finite planning horizon. Thus, there can be zero, finite or countably infinite number of adoptions. In case in which it is zero, no adoption occurs and the agent sticks to the initial technology in use leading to a technological sclerosis. In the case where it is finite, there exists at least one technology adoption but sticking occurs for some more advanced technology.

3 The Resolution: Multi-Stage Optimal Control

In order to solve this optimization problem, we use the approach of multi-stage optimal control with finite horizon that recalls Pontryagin's maximum principle from a dynamic programming perspective, where the switching times appear in both the objective function and the state equation.

[6] Jovanovic and Rob (1997) and Mateos-Planas (2000) also assume the adoption of the frontier technology. In Parente (2000), the adoption requires a time-to-built investment and how expert a firm is in the operation a new technology depends on how close the adopted technology is to the frontier. Thus, it does not always predict the adoption of the frontier technology.

As a solution must determine also the optimal number of adoptions that will take place within a given planning horizon together with their timing as well, our resolution process is constituted from two parts. In the first part, taking an arbitrary number of technology upgrades J and assuming that the economy did the last switch at t_J; we solve the sub-problem starting at t_J concerning the final stage of our multi-stage optimization problem. Then, we proceed backwards taking into account the optimal paths and the optimal value functions that we obtained for each subsequent stage of our problem. Every adoption timing will be chosen taking into account that subsequent timing of adoptions will be decided optimally given the remaining time-span. For a given planning horizon, the arbitrary number of adoptions may not be consistent with the optimal choice. The second part of our resolution procedure will be then to find the optimal J. Let $V(T \mid n)$ denote the maximum value of the objective function, namely the welfare of the economy, conditional on the plan containing exactly n adoptions with a given planning horizon. Then the optimal J will be the solution to

$$J^* = \arg\max_n \{V(T \mid n) : n = 1, 2, 3, \ldots\}, \tag{9}$$

so that $V^* = V(T \mid J^*)$ will be the optimal welfare of our economy.[7]

3.1 Auxiliary Problem (J)

In order to give more insight for the resolution process of the multi-stage optimal control problem, we attempt to decompose the original multi-stage problem into a sequence of almost conventional problems. Then the Maximum Principle will be used to obtain a set of optimality conditions. We start with an auxiliary problem corresponding to the final stage of the multi-stage problem:

$$\max_{\{c(t)\}} \int_{t_J}^{T} u(c(t)) e^{-\rho t} dt$$

subject to

$$c(t) + I(t) = \left(1 - e^{-\theta(t - t_J)}\right) K(t); \tag{10}$$

$$\dot{K}(t) = q(t_J)\, I(t); \quad t_J \le t \le T; \tag{11}$$

$$K(t_J) = K_J \text{ and } K(T) \text{ free}. \tag{12}$$

Note that the starting time t_J and the starting point K_J for this auxiliary problem are considered to be exogenous. This implies that t_J is merely a given constant and the appearance of it in the integrand of the utility function doesn't cause any conflict when Pontryagin's Maximum Principle is applied. In this

[7] Asteriks (*) is used to denote optimal quantities.

problem, a control c is said to be admissible whenever there exists a corresponding solution for the state equation (11), that satisfies the initial and the final conditions (12), and c satisfies the control constraint (10). Let the Hamiltonian of this auxiliary problem for the final stage H_J be defined by

$$H_J(K(t), c(t), \lambda_J(t), t, t_j) = -\ln[c(t)]e^{-\rho t} + \lambda_J(t)\left[q(t_J)\left(1 - e^{-\theta(t-t_J)}\right)K(t) - c(t)\right], \quad (13)$$

where we assume a logarithmic utility function and λ represents the costate variable. Remember also that $q(t) = q(t_J) = e^{\gamma t_J}$, for $t \in [t_J, T]$. Referring to a standard result[8], the optimality conditions for this auxiliary problem can be summarized by the following lemma, with the shorthand notation as follows:

$$H_j^* = H_j(K^*(t), c^*(t), \lambda^*(t), t, t_j),$$
$$H_j|_s = H_j(K(s), c(s), \lambda(s), s, t_j).$$

Lemma 1 Let c^* be an optimal control function for auxiliary problem (J). Then it is necessary that there exists a costate λ^* such that K^* and λ^* satisfy the canonical equations together with the end-point condition, and c^* minimizes the Hamiltonian:

$$\min_c H_J(K^*(t), c, \lambda_J^*(t), t) = H_J^* \text{ almost everywhere on } [t_J, T], \quad (14)$$

$$\dot{K}^*(t) = \frac{\partial H_J^*}{\partial \lambda_J} \text{ and } \dot{\lambda}_J^*(t) = -\frac{\partial H_J^*}{\partial K} \quad (15)$$

$$\lambda_J^*(T) \leq 0 \text{ and } \lim_{t \to T} K^*(T)\lambda_J^*(T) = 0. \quad (16)$$

From the resolution of these well known conditions, since the optimal path of consumption clearly depends on t_J and K_J, we have $V_J^* = V_J^*(K_J, t_J)$ as the optimal value function of this problem, i.e.,

$$V_J^*(K_J, t_J) = \int_{t_J} \ln[c^*(K_J, t_J)]e^{-\rho t}dt. \quad (17)$$

Using $V_J^*(K_J, t_J)$ as defined above and following a dynamic programing principle, we move to the problem concerning the stage that incorporates the $J-1$'th generation of adopted technology.

3.2 Auxiliary Problem (J-1)

$$\max_{\{c(t), t_J\}} \int_{t_{J-1}}^{t_J} u(c(t))e^{-\rho t}dt + V_J^*(K_J, t_J)$$

[8] See for instance, Athans and Falb (1966), Tomiyama (1985) and Kamien and Schwarz (1991).

subject to

$$c(t)+I(t)=\left(1-e^{-\theta(t-t_{J-1})}\right)K(t); \tag{18}$$

$$\dot{K}(t)=q(t_{J-1})\,I(t);\quad t_{J-1}\le t<t_J; \tag{19}$$

$$K(t_{J-1})=K_{J-1}\ \text{and}\ K(t_J)\ \text{free}. \tag{20}$$

Assuming that t_J^* is an interior point in $[t_{J-1};T]$ makes the constraint $t_{J-1}\le t_J\le T$ inactive and we are left with an auxiliary problem with free end point and free terminal time. However, the problem is still not standard in the sense that t_J is a choice variable and it appears at the upper limit of integration, at the integrand through the evolution process of capital and resource constraint and at the optimal value function of J'th stage. The explicit dependence of V_J^* on t_J must be considered carefully. In this problem, t_J and $(c(t);K(t))$; $t\in[t_J;T]$ are said to be admissible whenever $t_J\in(t_{J-1};T)$ and whenever there exists a corresponding solution for the state equation that satisfies the initial and the final conditions and c satisfies the control constraint. The optimality conditions of this auxiliary problem are summarized in the following lemma.

Lemma 2 Let c^* and t_J^* be optimal for the auxiliary problem (J-1). Then there exists a costate variable λ_{J-1}^* such that

$$\dot{K}^*(t)=\frac{\partial H_{J-1}^*}{\partial\lambda_{J-1}}\ \text{and}\ \dot{\lambda}_{J-1}^*(t)=-\frac{\partial H_J^*}{\partial K}; \tag{21}$$

$$\min_c H_{J-1}(K^*(t);c;\lambda^*(t);t)=H_{J-1}^*\ \text{almost everywhere on}\ [t_{J-1};t_J^*); \tag{22}$$

and K^*; λ_{J-1}^* and H_{J-1}^* satisfy the following matching conditions that ensure the continuity and optimality of the problem at $t=t_J^*$:

$$K^*(t_J^*-)=K^*(t_J^*+); \tag{23}$$

$$\lambda_{J-1}^*(t_J^*)=-\frac{\partial V_J^*(K_J;t_J^*)}{\partial K_J}; \tag{24}$$

$$H_{J-1}^*\big|_{t_J^*}+\int_{t_{J-1}}^{t_J^*}\left(\frac{\partial H_{J-1}^*}{\partial t_J}\right)dt=\frac{\partial V_J^*(K_J;t_J^*)}{\partial t_J}. \tag{25}$$

Proof. The derivation of the optimality conditions with a variational view of the control problem are left to Appendix. ∎

Remark 1 Under the condition that V_J^* is twice continuously differentiable in K_J and t_J, the matching conditions can be reformulated by using the dynamic programming approach as follows:

$$\frac{\partial V_J^*(K_J;t_J^*)}{\partial K_J}=-\lambda_J^*(t_J^*); \tag{26}$$

$$\frac{\partial V_J^*(K_J;t_J^*)}{\partial t_J}=H_J^*\big|_{t_J^*}-\int_{t_J}\left(\frac{\partial H_J^*}{\partial t_J}\right)dt. \tag{27}$$

Therefore we have

$$\lambda_{J-1}^*(t_J^*) = \lambda_J^*(t_J^*)\ ; \tag{28}$$

$$H_J^*|_{t_J^*} - H_{J-1}^*|_{t_J^*} = \int_{t_{J-1}}^{t_J^*} \left(\frac{\partial H_{J-1}^*}{\partial t_J}\right) dt + \int_{t_J}^{\infty} \left(\frac{\partial H_J^*}{\partial t_J}\right) dt. \tag{29}$$

Proof. It follows directly from Tomiyama and Rossana (1989), using the dynamic programming technique and computing the partial derivative $\frac{\partial V_J^*(K_J, t_J^*)}{\partial t_J}$ by Leibnitz' rule. The details are in Appendix. ■

Note that the consecutive stages of our problem are connected through the co-state variable. We know that in general the value of the co-state variable at the initial time of the planning horizon represents the marginal valuation at the optimum, due to a marginal increase in the initial value of the state variable. Then the value of the co-state variable at the instant of switching to a new technology represents the rate of change of the optimal value function V_J^* with respect to the change in the capital stock. In other words, $\lambda_J^*(t_J^*) - \lambda_{J-1}^*(t_J^*)$ represents the marginal gain in welfare terms from a marginal increase in the capital stock at the finite switching instant.

Remark 2 In contrast with Tomiyama (1985) and Makris (2001), the Hamiltonians do not match at the switching time because of the fact that the state equation and the objective function depends on the switching time explicitly. If this dependence is removed, it is clear that the conditions would reduce to those of Tomiyama (1985) and Makris (2001).

The resolution of these conditions leads to the optimal value V_{J-1}^* for the auxiliary problem (J-1) defined as follows:

$$V_{J-1}^*(K_{J-1}; t_{J-1}; t_J^*) = \int_{t_{J-1}}^{t_J^*} \ln[c^*(K_{J-1}; t_{J-1})] e^{-\rho t} dt + V_J^*(K_J(K_{J-1}; t_{J-1}); t_J^*). \tag{30}$$

Recall that, when posing the auxiliary problem (J-1) we have taken the initial time and the initial level of capital stock, t_{J-1} and K_{J-1} respectively, exogenously given. Thus, the optimal time of switching to the J'th technology clearly depends on this initial time. t_J^* is assured to be the overall optimum of the $V_{J-1}^*(K_{J-1}; t_{J-1}; t_J^*)$ due to the matching condition (25) that reflects the marginal gain in welfare terms from extending the horizon in which the $J-1'$th technology in use to the detriment of the J'th technology one. If it is optimal to switch to the J'th technology in an interior and well-defined instant of time, then the marginal benefit from extending the $J-1'$th technology horizon to the detriment of the J'th one should be equal to the marginal switching cost at the instant of switching.

Remark 3 Under the degenerate cases, where the optimal switching is either at t_{J-1} or at T, the condition (25) should be replaced with

$$\text{at } t_J^* = t_{J-1}; \quad H_{J-1}^*\big|_{t_{J-1}} \geq \frac{\partial V_J^*(K^*(t_{J-1}); t_{J-1})}{\partial t_J}$$

and

$$\text{at } t_J^* = T; \quad H_{J-1}^*\big|_T \leq \frac{\partial V_J^*(K^*(T); T)}{\partial t_J}.$$

These can be easily shown by checking the variation of the value function with respect to a perturbation in t_J^* at both points.

Up to now, by means of Lemma (1) and Lemma (2), we have characterized the optimal time of switching to the J'th technology, t_J^* for a given t_{J-1}. In order to characterize the optimal timing of switching to the J-1'th technology, we have to study the auxiliary problem of the J-2'th stage. t_{J-1}^* will then be chosen to be the overall optimum of the $V_{J-2}^*\left(K_{J-2}; t_{J-2}, t_{J-1}^*, t_J^*\right)$ by means of matching conditions. Continuing to follow the same resolution process, by studying the auxiliary problem of each stage backwards $(j = J-2, J-3, \dots, 0)$, our original problem of multi-stage optimal control reduces to a almost conventional problem. The optimal timing of technology adoptions, $t_1^*, t_2^*, \dots, t_{J-1}^*, t_J^*$ will then be the solution of a simultaneous equations system, stemming from the matching conditions. To sum up, we have established the following theorem on the optimality conditions for handling the multi-stage optimal control problem concerned with the optimal pattern of technology adoptions.

Theorem 1 Suppose that $\left(c^*; t_1^*, t_2^*, \dots, t_{J-1}^*, t_J^*\right)$, $0 < t_1^* < t_2^* < \dots < t_J^* < T$, is an optimal path for our discounted multi-stage optimization problem with J switches where J is a finite real number. Then it is necessary that there exist $\lambda_j^*(t)$; $j = 0, 1, 2, \dots, J$ such that

i) $c^*(t)$ minimizes the Hamiltonians,

$$\min_c H_0\left(K^*(t); c; \lambda_0^*(t); t\right) = H_0^* \text{ a.e. on } [0, t_1) \tag{31}$$

$$\min_c H_j\left(K^*(t); c; \lambda_j^*(t); t\right) = H_j^* \text{ a.e. on } [t_j, t_{j+1});\ j = 1, 2, \dots, J-1 \tag{32}$$

$$\min_c H_J\left(K^*(t); c; \lambda_J^*(t); t\right) = H_J^* \text{ a.e. on } [t_J, T). \tag{33}$$

ii) canonical equations are satisfied by the optimal trajectory,

$$\dot{K}^*(t) = \frac{\partial H_j^*}{\partial \lambda_j}; \tag{34}$$

$$\dot{\lambda}_j^*(t) = -\frac{\partial H_j^*}{\partial K} \tag{35}$$

except at points of discontinuity of $c^*(t)$ together with the initial and terminal conditions

$$K(0) = K_0 \text{ and } \lim_{t \to T} K^*(T)\,\lambda_J^*(T) = 0. \tag{36}$$

iii) Finally, the following matching conditions are satisfied at the switching instants

$$K^*\left(t_j^{*-}\right) = K^*\left(t_j^{*+}\right); \quad j = 1,2,\ldots,J \tag{37}$$

$$\lambda_j^*\left(t_{j+1}^*\right) = -\frac{\partial V_{j+1}^*}{\partial K}; \quad j = 0,1,2,\ldots,J-1 \tag{38}$$

$$H_0^*|_{t_1^*} + \int_0^{t_1^*}\left(\frac{\partial H_0^*}{\partial t_1}\right)dt = \frac{\partial V_1^*}{\partial t_1}; \tag{39}$$

$$H_j^*|_{t_{j+1}^*} + \int_{t_j^*}^{t_{j+1}^*}\left(\frac{\partial H_j^*}{\partial t_{j+1}}\right)dt = \frac{\partial V_{j+1}^*}{\partial t_{j+1}}; \ ; j = 1,2,\ldots,J-1. \tag{40}$$

Remark 4 Under condition that V_j^* is twice continuously differentiable in K_j and t_j; $j = 1,2,\ldots,J$, the matching conditions (38), (39) and (40) can be modified with

$$\lambda_{j-1}^*\left(t_j^*\right) = \lambda_j^*\left(t_j^*\right) \tag{41}$$

$$H_j^*|_{t_j^*} - H_{j-1}^*|_{t_j^*} = \int_{t_{j-1}^*}^{t_j^*}\left(\frac{\partial H_{j-1}^*}{\partial t_j}\right)dt + \int_{t_j^*}^{t_{j+1}^*}\left(\frac{\partial H_j^*}{\partial t_j}\right)dt. \tag{42}$$

Remark 5 If t_j^*; $j = 1,2,\ldots,J$ were fixed so that they were not choice variables then the conditions (42) would have been unnecessary.

Proof. Follows directly from Lemma (1), Lemma (2) and the Remark (1). On the other hand, the conditions can also be obtained through the technique of Calculus of Variations.[9] The derivation of these conditions for a general class of multi-stage optimal control problem are left to the Appendix. ■

From the resolution of these necessary conditions, we obtain the optimal path of consumption for a given, J number of switches as follows:

$$c^*(t) = \frac{\rho K_0}{q(0)}\frac{e^{(q(t_0)-\rho)(t-t_0)}}{1-e^{-\rho T}}; \quad 0 \le t < t_1^*$$

$$c^*(t) = \frac{\rho K_j}{q(t_j)}\frac{1}{1-e^{-\rho(T-t_j)}}e^{(q(t_j)-\rho)(t-t_j)-\frac{q(t_j)}{\theta}\left(1-e^{-\theta(t-t_j)}\right)};$$

$$t_j \le t < t_{j+1}; \quad j = 1,2,\ldots,J$$

[9] See Makris (2001) for the derivation of the optimality conditions for the infinite horizon discounted two-stage optimal control problem. However, these conditions can not be applied for our purpose as the explicit switching point dependence of the objective function and the equation of motion is not considered.

where

$$K_j = K_{j-1}\left(\frac{1-e^{-\rho(T-t_j)}}{1-e^{-\rho(T-t_{j-1})}}\right)e^{(q(t_{j-1})-\rho)(t_j-t_{j-1})-\frac{q(t_{j-1})}{\theta}\left(1-e^{-\theta(t_j-t_{j-1})}\right)};$$

$$j = 2,3,\dots,J$$

$$K_1 = K_0\left(\frac{1-e^{-\rho(T-t_1)}}{1-e^{-\rho T}}\right)e^{(q(0)-\rho)t_1}.$$

Notice that the consumption path is discontinuous at t_j^*; $j = 1,2,\dots,J$. Since $q(t_j) > q(t_{j-1})$, we have the level of consumption along the old technology regime is higher than in the new regime in each of the technology switches. It is due to the fact that a higher level of embodied technical change is associated with a decrease in the relative price of capital, which increases the user cost of capital by the so called obsolescence costs. This is a well known property related to the embodied technical change since Solow (1960) and shown to be non-negligible.[10]

On the other hand, by switching to a new technology which is always the frontier technology in our set-up, the economy will experience an improvement in the efficiency of capital goods, referred to as a growth rate advantage. In this sense, the matching conditions given in (39) and (40) will solve the trade-off between the growth rate advantage and the associated costs due to obsolescence and the loss in expertise. In cases where the growth rate advantage will not be sufficient to compensate for the obsolescence and the learning costs associated to switching from an initial technology in use to a new one, the economy would face a technological sclerosis.

Having established the optimality conditions for the multi-stage optimal control problem for a given number of switches, now we turn our attention to the optimal number of switches within a given planning horizon. As we have noted before, the optimal number of switches will be determined as a solution to

$$J^* = \arg\max_n \{V(T \mid n) : n = 1,2,3,\dots\}. \tag{43}$$

As we are not able to derive the closed form representation of the optimal number of switches and thus, the optimal timing of technology adoptions, we resort to numerical analysis.[11]

4 Numerical Analysis

In this section, we illustrate the pattern of technology adoptions and how it responds to the changes in growth rate of technology (γ), the speed of learning

[10] See Krussel (1998) and Boucekkine, del Rio and Licandro (2000).

[11] See Boucekkine, Saglam and Valée (2002), in which an analytical computation of optimal timing of a single technology adoption is presented via two-stage optimal control.

(µ) and the discount factor (½). We study the role of finite planning horizon on the pattern of technology adoptions. Our concern is on how the optimal number of switches and thus the duration between two consecutive ones, namely the tenures of the technologies to be adopted, differ with respect to the changes in the planning horizon. We know that, if the planning horizon were infinite so that the optimal number of switches were infinite then a solution would have been consisted of equally spaced adoptions.[12] Our departure from this case leads us to obtain time varying durations between consecutive adoptions. These variations occur mainly at the initial and the final technologies in use. Due to finite planning horizon, the duration left after the last switch may not be sufficient enough to cover the costs of a further switch that leads to a higher duration for the final technology to be in use.

The parameters of the model that we will assign values are the discount rate ½; the exogenous rate of growth of the technology frontier °; the speed of learning µ and finally the planning horizon T. There is a vast number of studies trying to asses the value of the speed of learning in the literature.[13] Jovanovic and Nyarko (1995), shows that the increases in the productivity of individual plants following their start-ups are typically large and realized in a short period of time. Bahk and Gort (1993), using a panel data of new plants from 15 different industries, estimates that learning is realized within 6 years and at each year 1 % increase is associated to the increase in the output. Consistent with these, the speed of learning is calibrated as 0.7 in Mateos-Planas (1997). Accordingly, we set µ = 0.7. We assigned a value of 4% for °; trying to be as close as possible to Greenwood and Yorukoglu (1997) and Del Rio (2002). Greenwood and Yorukoglu (1997) find that the decline of the relative price of equipment that can be seen as a proxy of the rate of investment specific technological progress has changed from 3.3% to 4% after 1974 in U.S. economy. In a very recent work, Del Rio (2002) finds that investment specific technological progress amounts to 2.27 %. Finally, we arbitrarily set the planning horizon to 40 years and the discount rate to 0.02. The initial level of capital stock K_0 is normalized to 1. The table just below summarizes the benchmark parameterization we chose.

Table 1: Parametric values as a benchmark

°	½	µ	T	K_0
0.04	0.02	0.7	40	1

Given these parameter values, in case of a single switch the maximum welfare level that can been attained is $V^*(T|1) = 470.663$. Allowing for only two switches, we obtain that the resulting optimal welfare level is $V^*(T|2) = 496.698$. Three switches within 40 years time lead to the optimal welfare of $V^*(T|3) = 499.975$. The optimal pattern of technology switches are given in

[12] See Parente (1994, 2000) and Iacompetta (2001).

[13] For a general survey, see Jovanovic (1997) and Greenwood and Jovanovic (1998).

Table 2. Note that, up to three switches the optimal solution in each case features learning of each technology that will be adopted. However, when we check the optimal pattern of technology adoptions incorporating 4 or more switches, the solution exhibits that some technologies may be skipped without learning. In such cases, only the technologies that are to be adopted earlier together with the final technology will have enough respite to be learned. For instance, in case of 4 switches the optimal time of adoptions are $t_1^* = 9.72$; $t_2^* = 19.53$; $t_3^* = 29.84$ and $t_4^* = 31.97$. This leads to a welfare of $V^*(T|4) = 490.393$. Note that $V^*(T|4) < V^*(T|3)$. This actually means that if we check for the corner solution in case of 4 switches where $t_4^* = T$; we will end up with a higher level of welfare than $V^*(T|4)$. This is due to the fact that the resulting improvement in the efficiency of capital goods sector is not enough to cover the obsolescence cost and the loss in expertise for four or more number of switches within a given duration of planning horizon. Thus, for our benchmark setting, we conclude that the optimal adoption plan features three adoptions ($J^* = 3$).

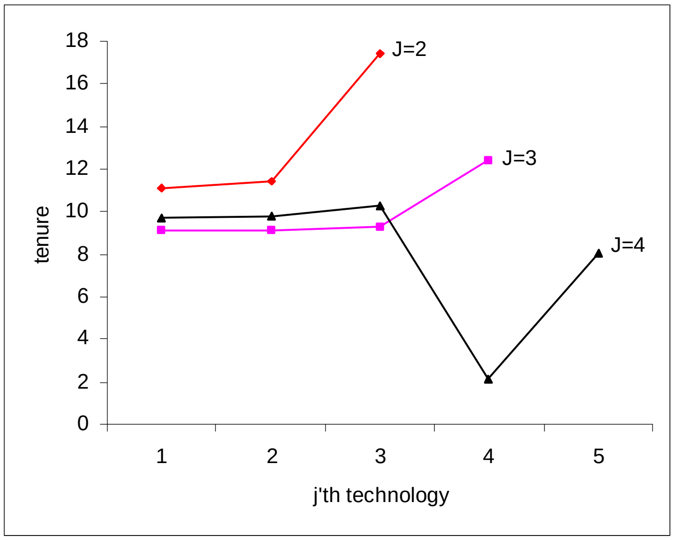
Table 2: Determination of the optimal number of adoptions

J	t_1^*	t_2^*	t_3^*	t_4^*	t_5^*	$V^*(T\|J)$
1	15.57					470.663
2	11.09	22.54				496.698
3	9.14	18.29	27.61			499.975
4	9.72	19.53	29.84	31.97		490.393
5	10.93	22.19	23.95	25.82	27.84	466.490

In contrast with Mateos-Planas (2000), the admissible adoption plans that incorporate switching to a better technology without gaining enough experience with the old one are not optimal. Our model leads to a solution that incorporates an optimal number of switches which lets each adopted technology to be learned. This is mainly due to the fact that the gain in the rate of embodied technological change is associated in our framework with a reduction in the relative price of capital. This gain is accompanied by an increase in the price of consumption goods with respect to the price of capital goods. This leads to a re-assignment of the resources of the economy towards the capital goods sector so that the consumption should drop from the date of each switch. As noted before, we refer to this cost as the obsolescence inherent to the embodiment.

The obsolescence cost together with the cost due to the loss in expertise of a technology switch put an upper bound on the number of switches within a finite planning horizon. It is clear from Table 2 that the optimal welfare shows a ¤ ¡ shaped relationship with respect to the number of switches. For a given duration of planning horizon, as the number of adoptions increases, the time required for the early adoptions to occur decreases until the optimal number of switches is reached and then it increases. This increase occurs because of the impossibility of learning each adopted technology within a given time horizon so that it is optimal to increase the tenure of the technologies which are endowed with more level of consumption.

Figure 1: The effect of changes in J on the pattern of tenures



For the benchmark values of the parameters, note that the optimal adoption plan exhibits almost a stationary path of tenures except than the final technology in use. Figure 1, illustrates this clearly. The tenures of the first three technologies in use have a slightly increasing path and are around 9.14 years, whereas the tenure of the final technology is 12.39 years. As the loss in expertise following a technology switch is not permanent and the economy is supposed to learn the new technology as time elapses, the duration left should be sufficient to asymptotically eliminate this expertise gap. However, the duration left after the last switch at $t_3^* = 27.61$ is not sufficient enough to cover these costs due to obsolescence and the loss in experience of a further switch. Thus, in contrast with Parente (1994, 2000), a non-stationarity in the durations of the adopted technologies to be in use arises due to a finite planning horizon.

Following from Boucekkine, Saglam and Valée (2002), we know that the optimal adoption delay in the interior solution case of a single technology switch problem, is naturally increasing with respect to the time horizon. It is proven that a longer time horizon gives more opportunities to take advantage of the growth rate advantage, so that delaying adoption more in order to take advantage of the consumption level advantage of the old regime is indeed optimal. In line with these, when we consider the effect of an increase in T that does not alter the optimal number of switches, leads to a delay in the adoption of the new technologies so that the tenures of each adopted technology increase. The effects of changes in T on the optimal number of switches and thus on the dates when these switches will occur are given in Table 3. Compared with the benchmark setting, increasing the time horizon to 45 years does not alter the optimal number of technology adoptions but the duration between two consecutive switches increases. Increasing further the time horizon to 50 years, increases the optimal number of switches and not surprisingly reduces the tenures of the technologies in use with respect to those in the case of $T = 45$. It must be also noted that, if the planning horizon is short enough ($T < 20$), the economy would stick to the initial technology in use and end up with a technological sclerosis.

Table 3: The effect of changes in T

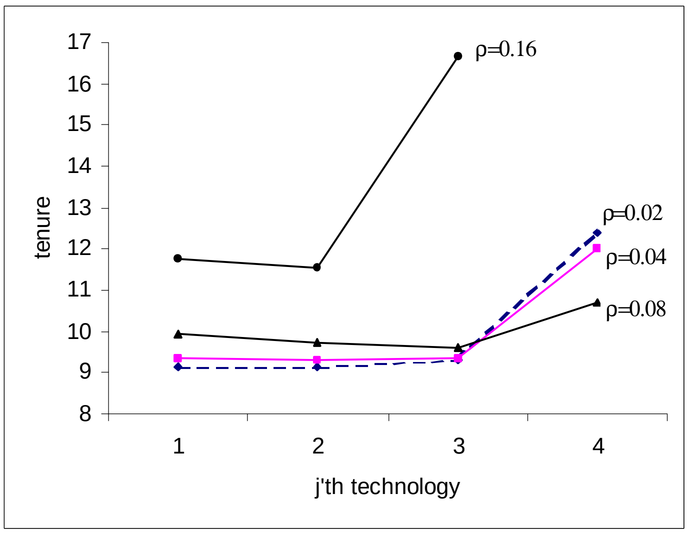
T	J^*	t_1^*	t_2^*	t_3^*	t_4^*	t_5^*	$V(T \mid J^*)$
20	1	8.95					106.062
25	1	10.15					175.144
30	2	9.06	18.18				261.849
35	2	9.94	20.15				370.300
40	3	9.14	18.29	27.61			499.975
45	3	9.91	19.86	30.13			653.180
50	4	9.21	18.42	27.69	37.18		832.021
55	4	9.93	19.82	29.80	40.11		1036.82
60	5	9.29	18.56	27.85	37.23	46.85	1272.71
65	5	9.98	19.89	29.79	39.78	50.11	1537.82

When we analyze the behavior of the adoption pattern with respect to time discounting, we observe that a higher impatience rate tends to delay adoption. In other words, it increases the duration between two consecutive technology adoptions. The tenures of the early adopted technologies tend to increase whereas the tenures of the late adopted technologies tend to decrease. Figure 2 illustrates this clearly. Even an economy that is impatient enough may observe a decrease in the optimal number of adoptions. The delay in adoption of the more advanced technologies as ρ increases is due to the increases in the obsolescence costs. However, as it is proved for a single switch case in Boucekkine, Saglam and Valée (2002), there may exist an opposing effect that tends to accelerate adoption since it would lead the economy to start to learn the new technology sooner and rise the growth rate advantage due to technology switching. In our set-up, with our parameterization the effect due to obsolescence costs dominate the effects due to growth rate advantage as ρ increases. This is because of the fact that we incorporate an exogenously growing technology frontier and technology specific learning by doing that tend to increase the effect of the obsolescence costs and the expertise loss in case of a technology switch. For our parameterization, the effect of the time discounting on the pattern of technology adoption is summarized in Table 4.

Table 4: The effect of changes in ρ

ρ	J^*	t_1^*	t_2^*	t_3^*	$V(T \mid J^*)$
0.01	3	9.06	18.16	27.46	665.838
0.02	3	9.14	18.29	27.61	499.975
0.04	3	9.34	18.63	27.99	288.495
0.08	3	9.96	19.69	29.28	106.465
0.16	2	11.78	23.33		22.340

Figure 2: The effect of changes in ρ on the pattern of tenures



In contrast with the effect of an increase in the discounting rate, an increase in the exogenous growth rate of the frontier technology leads to an acceleration in adopting the new technology. This is far from surprising because of the associated increase in the growth rate advantage that dominates the increase in the costs due to obsolescence and the lost experience. Figure 3 depicts the effect of a change in γ on the optimal pattern of technology adoptions. As γ increases, the duration between two consecutive adoptions will be reduced and accordingly the tenures of the adopted technologies except than the final technology will be lowered. A sufficient increase in γ would also increase the optimal number of adoptions within a given planning horizon. This leads to a further decrease in the duration between two consecutive adoptions. Table 5 reports the optimal number of switches and the optimal dates of switching for different values of technological progress.

Table 5: The effect of changes in γ

γ	J^*	t_1^*	t_2^*	t_3^*	t_4^*	t_5^*	$V(T \mid J^*)$
0.01	1	18.24					380.332
0.02	2	12.63	25.27				405.229
0.04	3	9.14	18.29	27.61			499.975
0.05	3	8.90	17.80	26.93			570.204
0.06	4	7.45	14.84	22.30	29.92		660.512
0.08	5	6.28	12.47	18.73	25.07	31.57	922.624

Finally a higher value for θ means a faster learning process, which rises the incentives to switch. Table 6 reports the optimal number of switches and the associated optimal dates of switching to new technologies for different values of θ. An increase in θ accelerates the adoption so that the duration between two consecutive adoptions would decrease. In other words, it tends to decrease the tenures of the early adopted technologies and increase the tenure of the final technology to be used. Not surprisingly, a sufficient increase in the speed of learning would lead to an increase in the optimal number of adoptions as the time duration required to compensate for the loss in expertise of a technology switching will be reduced. In such cases, the optimal solution features a further decrease in the duration between two consecutive technology adoptions and thus, a higher tenure for the final technology to be used. Figure 4 depicts these clearly. Apart from these, under our benchmark parameterization where $\gamma = 0.04$; $\rho = 0.02$ and $T = 40$; for sufficiently low values of speed of learning ($\theta < 0.2$), the economy never switches to a better technology and simply sticks to the initial technology facing a technological sclerosis.

Table 6: The effect of changes in θ

θ	J^*	t_1^*	t_2^*	t_3^*	t_4^*	t_5^*	$V(T \mid J^*)$
0.2	1	17.95					406.498
0.4	2	12.20	24.36				454.837
0.6	2	11.29	22.88				486.700
0.7	3	9.14	18.29	27.61			499.975
0.8	3	8.88	17.82	27.00			510.877
1.0	4	7.46	14.93	22.44	30.07		527.541
1.2	4	7.17	14.37	21.65	29.16		541.586
1.4	5	6.26	12.52	18.81	25.13	31.56	552.211
1.6	5	6.07	12.16	18.28	24.47	30.85	561.778

Figure 3: The e¤ect of changes in ° on the pattern of tenures

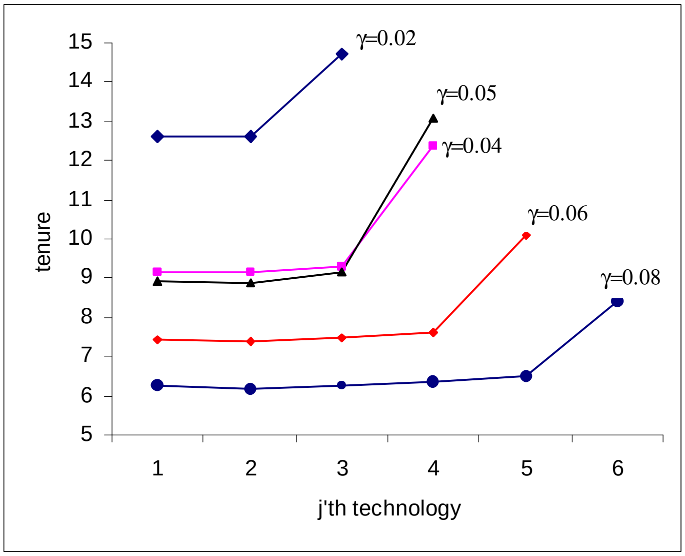
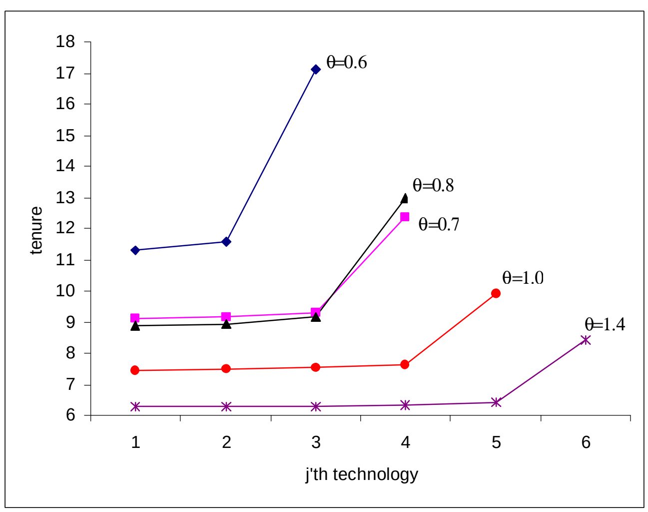


Figure 4: The e¤ect of changes in µ on the pattern of tenures



5 Conclusion

In this paper, we have applied multi-stage optimal control techniques to analyze the optimal pattern of technology adoption under embodiment when the planning horizon is finite. We have first presented the necessary and optimality conditions for multi-stage optimal control problem in which the adjustable switching times appear as an argument of the state equation. On the contrary to the existing literature, we have taken into account not only the learning costs associated to a technology upgrade but also the obsolescence costs inherent to embodiment. We have shown that the optimal number of switches and the timing of them depends on how the growth rate advantage compares to these learning and obsolescence costs. We have shown that the finite planning horizon induces a non-stationarity in the tenures of the adopted technologies. The non-stationarity arises mainly at the final technology in use as the duration left after the last switch may not be sufficient enough to cover the obsolescence and the learning costs of a further switch. We have mentioned that the optimal adoption plan involves sufficient tenures for each technology adopted in order to eliminate the expertise gaps that emerge after switches. That is to say, the adoption of technologies which would be skipped without gaining expertise is not optimal. We have provided numerically the effects of the planning horizon, growth rate of technology, discounting parameter and the speed of learning on the optimal pattern of technology adoptions.

We believe that this model provides a rich structure that could be used to examine several other types of optimal choice and timing problems. On growth-related issues, a decentralized equilibria could be considered in order to analyze the technological leapfrogging among firms and industries. It could be extended to analyze the effects of capital market efficiency, the role of uncertainty and the market structure on technology adoption. These are in our research agenda.

6 Appendix

We will consider a general form of multi-stage optimal control problem where the switching points appear at the integrand of the objective function and at the equation of motion explicitly when presenting the proofs.

$$\max_{\{c(t),\, t_1,\, t_2, \ldots, t_J\}} \int_0^{t_1} F_0(K(t), c(t), t_1, t)\,dt + \int_{t_1}^{t_2} F_1(K(t), c(t), t_1, t_2, t)\,dt + \ldots$$
$$\ldots + \int_{t_{J-1}}^{t_J} F_{J-1}(K(t), c(t), t_{J-1}, t_J, t)\,dt + \int_{t_J}^{T} F_J(K(t), c(t), t_J, t)\,dt$$

subject to the system equations,

$$\dot{K} = f_0(K(t);c(t);t_1;t);\ 0 \leq t < t_1; \tag{A.1}$$

$$\dot{K} = f_j(K(t);c(t);t_j;t_{j+1};t);\ t_j \leq t < t_{j+1};\ \forall j = 1;2;\ldots;J-1 \tag{A.2}$$

$$\dot{K} = f_J(K(t);c(t);t_J;t);\ t_J \leq t \leq T; \tag{A.3}$$

control constraints,

$$c(t) \in \Omega_0 \text{ a.e. on } 0 \leq t < t_1; \tag{A.4}$$

$$c(t) \in \Omega_j \text{ a.e. on } t_j \leq t < t_{j+1};\ \text{for } \forall j = 1;2;\ldots;J-1 \tag{A.5}$$

$$c(t) \in \Omega_J \text{ a.e. on } t_J \leq t \leq T \tag{A.6}$$

and the boundary conditions,

$$T \text{ given, } K(0) = K_0 \text{ given and } K(T) \text{ free,} \tag{A.7}$$

where F_j and $f_j;\ j = 0;1;\ldots;J$ are assumed to be at least continuously differentiable in the state variable K; time variable t and the switching points that they depend on and continuous in the control variable c: $c(t)$ is measurable on $[0;T]$ and its value is constrained to closed subsets of $1-$dimensional Euclidean space E^1; such that $\Omega \subset E^1$; presented in (A.4)-(A.6). $c(t)$ and $t_j;\ j = 1;2;\ldots;J$ are said to be admissible whenever $0 < t_1 < t_2 < \ldots < t_{J-1} < t_J < T$ and whenever there exists a corresponding solution for the state equation that satisfies the boundary conditions and c satisfies the control constraint.

Proof of Lemma 2. Auxiliary problem (J-1) can be written in the form of:

$$\max_{\{c;\,t_J\}} V_{J-1} = \int_{t_{J-1}}^{t_J} F_{J-1}(c;K;t;t_J)\,dt + V_J^*(K(t_J);t_J)$$

subject to

$$\dot{K}(t) = f_{J-1}(c;K;t;t_J);\ t_{J-1} \leq t < t_J; \tag{A.8}$$

$$K(t_{J-1}) = K_{J-1} \text{ and } K(t_J) \text{ free,} \tag{A.9}$$

Here c is assumed to be unconstrained so that c^* is an interior solution. Using the notion of Lagrange multipliers, we have:

$$\int_{t_{J-1}}^{t_J} \lambda_{J-1}(t)\left(f_{J-1}(c;K;t;t_J) - \dot{K}(t)\right)dt = 0: \tag{A.10}$$

Augmenting the objective function by this integral in (A.10) does not affect the solution. Thus, we can define the new objective function as

$$\Lambda_{J-1} \equiv \int_{t_{J-1}}^{t_J}\left[F_{J-1}(c;K;t;t_J) - \lambda_{J-1}(t)\left(f_{J-1}(c;K;t;t_J) - \dot{K}(t)\right)\right]dt$$
$$+ V_J^*(K(t_J);t_J): \tag{A.11}$$

The substitution of the Hamiltonian function that we have defined as

$$H_{J-1}(c;K;t;t_J;\lambda_{J-1}) = -F_{J-1}(c;K;t;t_J) + \lambda_{J-1}(t) f_{J-1}(c;K;t;t_J)$$

into (A.11), we obtain:

$$\Lambda_{J-1} = \int_{t_{J-1}}^{t_J} \left[-H_{J-1}(c;K;t;t_J;\lambda_{J-1}) - \int_{t_{J-1}}^{t_J} K(t)\dot{\lambda}_{J-1}(t) \right] dt$$
$$+\lambda_{J-1}(t_J)K(t_J) - \lambda_{J-1}(t_{J-1})K(t_{J-1}) + V_J^*(K(t_J);t_J): \quad \text{(A.12)}$$

The value of Λ_{J-1} depends on the time paths chosen for the variables, $c; K; \lambda_{J-1}$ and the values chosen for t_J as well. Let $(c^*(t);K^*(t))$ be an optimal path for every $t \in [t_{J-1};t_J)$ and t_J^* is an interior point in $[t_{J-1};T]$: If we perturb the $c^*(t)$ path by a perturbing curve $u(t)$; we can generate admissible "neighboring" control paths, $c(t) = c^*(t) + \varepsilon u(t)$; one for each value of ε: According to the equation of motion, there will then a corresponding perturbation in the $K^*(t)$ path that can be written as $K(t) = K^*(t) + \varepsilon x(t)$ together with $t_J = t_J^* + \varepsilon\Delta t_J$ and $K(t_J) = K(t_J^*) + \varepsilon\Delta K(t_J)$: Note that these imply $\frac{dt_J}{d\varepsilon} = \Delta t_J$ and $\frac{dK(t_J)}{d\varepsilon} = \Delta K(t_J)$: Now we can express Λ_{J-1} in terms of ε and apply the first-order condition $\frac{d\Lambda_{J-1}}{d\epsilon}|_{\varepsilon\to 0} = 0$; as follows:

$$\frac{d\Lambda_{J-1}}{d\epsilon} = \int_{t_{J-1}}^{t_J(\varepsilon)} \left\{ -\left(\frac{\partial H_{J-1}}{\partial c}u(t)\right) + \left(\frac{\partial H_{J-1}}{\partial K} + \dot{\lambda}_{J-1}(t)\right) x(t) \right\} dt$$
$$+ \left(\lambda_{J-1}(t_J) + \frac{\partial V_J^*(K(t_J);t_J)}{\partial K}\right) \Delta K(t_J)$$
$$+ \left(\frac{\partial V_J^*(K(t_J);t_J)}{\partial t_J} - H_{J-1}|_{t_J} - \int_{t_{J-1}}^{t_J(\varepsilon)} \frac{\partial H_{J-1}}{\partial t_J} \right) \Delta t_J = 0: \quad \text{(A.13)}$$

In (A.13), the integral contains arbitrary perturbing curves $u(t)$ and $x(t)$, and the other two components of the equation involve arbitrary $\Delta K(t_J)$ and Δt_J: In order to have (A.13) satisfied, each component of this derivative should be set equal to zero. By setting the integral component equal to zero, we have the following two conditions:

$$\dot{\lambda}^*_{J-1}(t) = -\frac{\partial H^*_{J-1}}{\partial K} \quad \text{(A.14)}$$

$$\frac{\partial H^*_{J-1}}{\partial c} = 0 \quad \text{(A.15)}$$

Equation (A.14) together with $\dot{K}^*(t) = \frac{\partial H^*_{J-1}}{\partial \lambda_{J-1}}$, $\forall t \in [t_{J-1};t_J)$ that we deduce from (A.8) constitutes the canonical equations. Equation (A.15), represents a weaker version of the $\min_c H_{J-1}$ when it is predicated on the assumption that H_{J-1} is differentiable with respect to c and there is an interior solution. Setting the two other components of the derivative (A.13) equal to zero and noting the continuity of the state variable leads to the matching conditions. $\blacksquare$

Proof of Remark 1.

$$
\begin{aligned}
\frac{dV_J^*(K(t_J);t_J)}{dt_J} &= \frac{d}{dt_J}\int_{t_J} F_J(c^*;K^*;t;t_J)\,dt \\
&= \frac{d}{dt_J}\int_{t_J} [F_J(c^*;K^*;t;t_J)+\lambda_J^*(t)(\dot{K}(t)-f_J(c;K;t;t_J))]dt \\
&= -F_J^*|_{t_J}+\int_{t_J}\left[\frac{\partial F_J^*}{\partial t_J}-\lambda_J^*(t)\left(\frac{\partial f_J^*}{\partial t_J}\right)\right]dt \\
&= -F_J^*|_{t_J}-\int_{t_J}\left(\frac{\partial H_J^*}{\partial t_J}\right)dt. \qquad (A.16)
\end{aligned}
$$

Apart from this,

$$
\begin{aligned}
\frac{dV_J^*(K(t_J);t_J)}{dt_J} &= \left(\frac{\partial V_J^*(K(t_J);t_J)}{\partial K(t_J)}\frac{dK(t_J)}{dt_J}\right)+\frac{\partial V_J^*(K(t_J);t_J)}{\partial t_J} \\
&= -\lambda_J^*(t_J)f_J^*|_{t_J}+\frac{\partial V_J^*(K(t_J);t_J)}{\partial t_J} \qquad (A.17)
\end{aligned}
$$

Therefore, combining the two equations (A.16) and (A.17) ends the proof. $\blacksquare$

Proof of Theorem 1. The set of canonical conditions, (34) and (35), together with the initial and the terminal conditions (36) and the minimization of the Hamiltonians, (31), (32) and (33) are direct consequences of Lemmas (1) and (2) due to the application of well-known Pontryagin's Maximum Principle to the auxiliary problem of the each stage. Under the condition that V_j^* is twice continuously differentiable in K_j and t_j; $j=1;2;:::;J$ the replacement of the matching conditions (38), (39) and (40) with those in (41) and (42) follows from Remark 1. The derivation of these conditions through the technique of Calculus of Variations is straightforward but rather lengthy and therefore we will concentrate on only the matching conditions.

Under the condition that V_j^* is twice continuously differentiable in K_j and t_j; $j=1;2;:::;J$; derive the first order variation, δD; for problem (G): After standard calculations, we obtain

$$
\begin{aligned}
\delta D = \sum_{j\in\{1,2,:::,J\}} &\left[H_j^*|_{t_j^*}-H_{j-1}^*|_{t_j^*}-\int_{t_{j-1}^*}^{t_j^*}\left(\frac{\partial H_{j-1}^*}{\partial t_j}\right)dt-\int_{t_j^*}^{t_{j+1}}\left(\frac{\partial H_j^*}{\partial t_j}\right)dt\right]\delta t_j \\
&+\left[\lambda_j^*\left(t_j^*\right)-\lambda_{j-1}^*\left(t_j^*\right)\right]\delta K_j; \qquad (A.18)
\end{aligned}
$$

in which δt_j and δK_j are any admissible perturbations in the switching instant t_j^* and in $K^*(t_j^*)$ respectively. $\delta D \le 0$ must be satisfied for any admissible path that is close to the optimal one. Consider the path that satisfies $\delta t_j = 0$; $j = 1, 2, ..., J$ then it is clear that we must have the condition (41) as a necessary one. On the other hand, suppose that the admissible paths satisfy $\delta K_j = 0$; $j = 1, 2, ..., J$. Then, given that $0 < t_1 < t_2 < ... < t_{J-1} < t_J < T$; we have the condition (42). Finally, if t_j^*; $j = 1, 2, ..., J$ were fixed so that they were not choice variables, we have that $\delta t_j = 0$; $j = 1, 2, ..., J$. So, Remark (5) follows directly. ¥

7 References

Athans, M. and P.L. Falb (1966), "Optimal Control", MacGraw-Hill, New York.

Bahk, B., and M. Gort (1993), "Decomposing Learning By Doing in New Plants", Journal of Political Economy 101, 561-583.

Boucekkine, R., F. del Río and O. Licandro (2000), "The importance of the Embodied Question Revisited", DP 9926, IRES, Université catholique de Louvain.

Boucekkine, R., C. Saglam and T. Vallée (2002), "Technology Adoption under Embodiment: A two-stage optimal control approach", Mimeo, IRES, Université catholique de Louvain.

Chiang, A. C. (1992), "Elements of Dynamic Optimization", McGraw-Hill, New York.

del Río F. (2002), "Neutral, Investment-Specific Technical Progress and the Productivity Slowdown", Recherches économiques de Louvain, 68 (1-2), 37-55.

Greenwood, J., Z. Hercowitz and P. Krusell (1997), "Long-Run Implications of Investment-Specific Technological Change," American Economic Review 87, 342-362.

Greenwood, J. and M. Yorukoglu (1997), "1974", Carnegie-Rochester Series on Public Policies 46, 49-95.

Greenwood, J. and B. Jovanovic (1998), "Accounting for Growth," in: New Directions in Productivity Analysis, E.Dean, M. Harper and C. Hulten (Eds), NBER.

Iacompetta, M. (2001), "Dissemination of Technology in Market and Planned Economies", Mimeo, New York University.

Jovanovic, B. and Y. Nyarko (1996), " Learning by Doing and the Choice of Technology", Econometrica 64, 1299-1310.

Jovanovic, B. (1997), "Learning and Growth", in D. Kreps and K. Wallis, eds., Advances in economics, Vol. 2 (Cambridge University Press, London): 318-339.

Jovanovic, B. and R. Rob (1997), "Solow vs. Solow: Machine Prices and Development", Mimeo, New York University.

Kamien, M.I. and N. L. Schwartz (1991), "Dynamic Optimization", North-Holland, Amsterdam.

Krusell, P. (1998), "Investment-Specific R&D and the Decline in the Relative Price of Capital," Journal of Economic Growth 3, 131-141.

Maccini, L.J. (1973), "On Optimal Delivery Lags", Journal of Economic Theory 6, 107-125.

Makris, M. (2001), "Necessary Conditions for Infinite Horizon Discounted Two-Stage Optimal Control Problems", Journal of Economic Dynamics and Control 25, 1935-1950.

Mateos-Planas, X. (2000), "Technology Adoption with Finite Horizons", DP 33, University of Southampton.

Mateos-Planas, X. (2001), "Schooling and Distortions in a Vintage Capital Model", Review of Economic Dynamics 4, 127-158.

Parente, S. (1994), "Technology Adoption, Learning by Doing, and Economic Growth", Journal of Economic Theory 63, 346-369.

Rossana, R. (1985), "Delivery Lags and Buffer Stocks in the Theory of Investment by the Firm", Journal of Economic Dynamics and Control 9, 153-193.

Tomiyama, K. (1985), "Two-Stage Optimal Control Problems and Optimality Conditions", Journal of Economic Dynamics and Control 9, 317-337.

Tomiyama K. and R. Rossana (1989), "Two-Stage Optimal Control Problems with an Explicit Switch Point Dependence: Optimality Criteria and an Example of Delivery Lags and Investment", Journal of Economic Dynamics and Control 13, 319-337.

Solow, R. (1960), "Investment and Technological Progress" in Kenneth J. Arrow, Samuel Karlin and Patrick Suppes, Eds., Mathematical Methods in the Social Sciences 1959, Stanford University Press, 89-104.