

## MODELLING VINTAGE STRUCTURES WITH DDEs: PRINCIPLES AND APPLICATIONS\*

---

*Raouf Boucekkine*<sup>†</sup>  
*IRES and CORE Université catholique de Louvain, Belgium*

*David de la Croix*<sup>‡</sup>  
*FNRS, IRES and CORE, Belgium*

*Omar Licandro*<sup>§</sup>  
*European University Institute, Florence, Italy*

*A comprehensive study of the linkages between demographic and economic variables should not only account for vintage specificity but also incorporate the relevant economic and demographic decisions in a complete optimal control set-up. A methodological set-up allowing to reach these objectives is described. In this framework, time is continuous but agents take discrete timing decisions. The mixture of continuous and discrete time yields differential-difference equations (DDEs). It is clearly shown that the approach allows for a relatively complete and rigorous analytical exploration in some special cases (mainly linear or quasi linear models), and for an easy computational appraisal in the general case.*

*Keywords: demography; economic growth; vintage structures; optimal control; differential-difference equations; state-dependence*

\*This article is an extended version of a paper prepared for the October 2003 Viennese Vintage Workshop. Special thanks to Gustav Feichtinger and Valdimir Veliov for suggesting and discussing this paper, and to two anonymous referees for useful comments. Boucekkine and de la Croix acknowledge the support of the Belgian research programs “Poles d’Attraction inter-universitaires” PAI P5/21, and “Action de Recherches Concertée” ARC 99/04-234 and ARC 03/08-302. Licandro acknowledges the support of the Spanish Ministry of Sciences and Technology (project SEC2000-0260).

<sup>†</sup>Corresponding author. IRES and CORE, Université catholique de Louvain, Place Montesquieu, 3, B-1348 Louvain-la-Neuve (Belgique). boucekkine@ires.ucl.ac.be

<sup>‡</sup>FNRS, IRES and CORE. delacroix@ires.ucl.ac.be

<sup>§</sup>European University Institute, Firenze, and FEDEA, Madrid. Omar.licandro@IUE.it

## 1 INTRODUCTION

The analysis of the relationship between demographic trends and economic growth has recently regained interest. Departing from the typical *ad hoc* treatment of demographic variables in the neoclassical models, many ongoing research programs are now attempting to deeply study the channels through which the demographic variables interact with economic growth. Besides the fundamental problem of the demographic transition (see, for example, Galor and Weil, 1999), the main issues concern the observed nonlinear relationships between economic growth and some demographic variables, as documented by Kelley and Schmidt (1995). Among their empirical findings, they identify an ambiguous effect of crude death rates. It appears that a decrease in the crude death rate increases economic growth, especially in the least-developing countries, where mortality reduction is concentrated in the younger and working ages. This is less true in developed countries because this reduction mainly affects the retired cohorts. Roughly speaking, economic growth is slowed by the deaths of the workers but can be enhanced by the deaths of dependents.

Consequently, the modelling of the age structure becomes indispensable to account for the demographic stylized facts of economic growth. When considering such a structure, the modeler has to specify the determinants of cohorts' differences, in particular those determinants concerning human capital and labor force participation. In de la Croix and Licandro (1999), and Boucekkine, de la Croix and Licandro (2002), a key vintage specificity element is that different generations have different education experiences. As a consequence, the aggregate stock of human capital, built from the human capital of the different cohorts and their participation decisions, depends on this vintage structure. Demographic decisions such like schooling and retirement decisions are endogenous and affect the vintage structure of the labor force.

Therefore, a full description of the linkages between demographic and economic variables should not only account for vintage specificity but also incorporate the relevant economic and demographic decisions in a complete optimal control set-up. These are the objectives pursued in the approach we have been using first on vintage capital growth models (see, for example, Boucekkine, Germain and Licandro, 1997), then on demographic vintage capital models (Boucekkine, de la Croix and Licandro, 2002). In our framework, time is continuous but the agents take discrete timing decisions (as schooling time or machines' scrapping time). The mixture of continuous and discrete time yields differential-difference equations (DDEs). Hereafter, we shall refer to this approach as the DDE approach.

The use of DDEs in economic dynamics traces back to Kalecki (1935) and Leontief (1961). However, in both papers, there is no optimization. The first optimal control-based model with vintage capital is the firm problem treated by Malcomson (1975). An application to optimal growth theory has been provided recently by Benhabib and Rustichini (1991). De la Croix and Licandro (1999) have provided the first application to the field of economic growth and demography.

This paper is devoted to give the main principles of the DDE approach with several applications, in particular to economic demography, in order to clarify its contributions and its scope. Naturally, the vintage structure can be modelled following alternative approaches, among which partial differential equations (PDE) are the most natural (see Barucci and Gozzi, 2001, and Feichtinger et al., forthcoming). In particular, the PDE modelling has the advantage to allow for the incorporation of learning effects, which comes from the fact that within this framework, investing in non-frontier technologies is possible. In contrast, as it will be clear in this paper, the DDE approach does not allow for this possibility. On the other hand, typical PDE modelling of the vintage capital structure takes as given the maximal age of capital, while this magnitude is an economic decision in the DDE approach. Therefore, both approaches have their relative advantages. This paper describes in details the DDE approach and shows how it can be useful for mathematical demography modelers.

This paper is organized as follows. The next section gives the general principles of the DDE approach and summarizes the early applications. In particular, it shows that this approach is perfectly appropriate to model creative destruction phenomena, such as the optimal scrapping time of obsolete technologies and capital goods and their replacement by more advanced and efficient ones. Section 3 is the demographic contribution of the paper. It builds on the vintage human capital model developed by Boucekkine, de la Croix and Licandro (2002) to address further demographic issues (such like the consequences of a drop in fertility) using some adequately calibrated versions of this model. Section 4 reviews some of the mathematical aspects of the methodology and makes clear the limits of the DDE approach, in particular when optimal control is involved.

## **2 THE DDE APPROACH: GENERAL SPECIFICATIONS, EXAMPLES AND EARLY APPLICATIONS**

This section explains our approach in the most general formulation so as to clearly illustrate its principles and to make clear its scope. Examples are given along the way.

## 2.1 Modelling the Age Structure

Time is continuous, and it is represented by  $t$ ,  $-\infty \leq t \leq \infty$ . Machines produced at time  $v$  belong to the same cohort, the so-called “vintage  $v$ ”. We denote by  $x(v)$ , the amount of variable  $x$  of vintage  $v$  at  $t$ . Therefore, at any date  $t$ , there exists a distribution of items  $x$  indexed by their vintage  $v$ . The aggregation of this age-structured population is a key issue. To get an immediate idea about it, think of  $x(v)$  as the number of machines of generation  $v$ . Generally, computing the aggregate active stock of capital at  $t$  is not the simple addition of all operated machines at  $t$  whatever their vintages. One should typically account for the vintage-specific characteristics of the machines: for example, capital depreciation is necessarily larger for old capital goods. Moreover, the age structure itself may be altered by the decision-maker: for many good reasons, the firms may decide to scrap all the machines older than a certain age. To cope with all these possible ingredients, let us adopt the following general formulation for the aggregate stock at  $t$ , say  $X(t)$ :

$$X(t) = \int_{t-D_2(t)}^{t-D_1(t)} x(v)\psi(v, t) dv, \quad (1)$$

where:

- (i)  $0 \leq D_1(t) < D_2(t) \leq \infty$ , are the *timing* or *delay* variables, which are generally control variables,
- (ii)  $\psi(v, t)$  represents physical depreciation, survival probability and/or any vintage-specific weight.

The following examples make clear the scope of this formulation.

### **Example 1:** Vintage capital growth models

For any date  $t$ , denote by  $i(v)$  the amount of capital of vintage  $v$ ,  $v \leq t$ . Also assume that due to technological progress, the new generations of capital are increasingly more efficient than the old ones, so that the firms typically decide to scrap the oldest and least efficient machines. Denote by  $T(t)$  the scrapping age. The active capital stock can be written as;

$$K(t) = \int_{t-T(t)}^t i(v)e^{-\delta(t-v)} dv, \quad (2)$$

with:  $\delta \geq 0$ , the capital depreciation rate, and  $T(t)$ , the oldest capital goods still in use at  $t$ . We are in the case:  $\psi(v, t) = e^{-\delta(t-v)}$ ,  $D_1(t) = 0$  and

$D_2(t) = T(t)$ . Typically,  $T(t)$  is a control variable and will be treated as such later in this section.

**Example 2:** Vintage capital with time to build

As a simple extension to our canonical Example 1, suppose that at time  $t$ , it takes  $B(t)$  units of time for any vintage capital to be productive (time to build). This specification is in line with Kalecki's lag (1935). In such a case, the active capital stock looks like:

$$K(t) = \int_{t-T(t)}^{t-B(t)} i(v)e^{-\delta(t-v)} dv, \quad (3)$$

where  $0 < B(t) < T(t) : D_1(t) = B(t)$  and  $D_2(t) = T(t)$ .

**Example 3:** Demographic models with an explicit age structure

At any time  $t$ , denote by  $h(v)$  the human capital of the cohort (or generation) born at  $v$ ,  $v \leq t$ . Suppose that  $t - T(t)$  is the last generation that entered the job market at  $t$ , because of time spent at school by all individuals,  $T(t)$ . If  $t - A(t)$  is the last generation still at work, where  $A(t)$  could be the maximal age attainable, then the aggregate stock of human capital available at time  $t$  is:

$$H(t) = \int_{t-A(t)}^{t-T(t)} h(v)e^{nv} m(t-v) dv, \quad (4)$$

where:  $n$  is the growth rate of population,  $e^{nv}$  is size of the cohort born at  $v$ , and  $m(t-v)$  is the probability for an individual born at  $v$  to be still alive at  $t$ . In this example,  $\psi(v, t) = e^{nv} m(t-v)$ ,  $D_1(t) = T(t)$  and  $D_2(t) = A(t)$ , with  $T(t) < A(t)$ . As will be made clear in Section 3,  $T(t)$  may be a control variable, while  $A(t)$  is typically an outcome of the postulated survival law.

**Example 4:** Incorporating retirement into the demographic models

A more realistic demographic model must take into account that generally individuals do not work until death: retirement is one of their most salient decisions. In our framework, retirement is just a timing and control variable more. Indeed, let  $P(t)$  be the last generation that retired at  $t$ , the aggregate stock of human capital available at time  $t$  becomes:

$$H(t) = \int_{t-P(t)}^{t-T(t)} h(v)e^{nv} m(t-v) dv, \quad (5)$$

In this case,  $D_1(t) = T(t)$  and  $D_2(t) = P(t)$ . Example 4 is studied in detail in Section 3, once complemented with an optimal control devices for the timing variables  $T(t)$  and  $P(t)$ . The next sub-section is devoted to show how the DDEs are derived from the above integral items, and why such a differentiation, though mathematically unnecessary, allows to better capturing and representing the economic mechanisms at work in the model.

## 2.2 The Induced DDE Structure and its Economic Interpretation

To unburden the presentation, consider the general Equation (1) with  $\psi(v, t) = 1, \forall v, t$ , namely:

$$X(t) = \int_{t-D_2(t)}^{t-D_1(t)} x(v)dv. \quad (6)$$

Assume that:

- (i)  $x(v)$  is piecewise continuous on  $R$ ,
- (ii)  $D_i(t), i = 1, 2$ , are piecewise differentiable on  $R$ .

Then, one can obtain by differentiation of (1):

$$X'(t) = x(t - D_1(t))(1 - D_1'(t)) - x(t - D_2(t))(1 - D_2'(t)). \quad (7)$$

Notice that the obtained differential equation includes lagged and possibly state-dependent terms that feature the DDE ingredient, and makes the mathematical problem so specific. Obviously, the lags are there because they first appear in the original integral equation. Moreover, as one may suspect, it is generally possible to tackle directly the original delayed-integral equation. We come back to this specific issue in the mathematical Section 4. Nonetheless, we proceed by differentiation in order to shed a closer light on the mechanisms at work in the models, which turns out to be much less easy on the integral form (1). Hereafter, we shall refer to the DDE (7) as a *creation-destruction motion*. We argue that such motions are extremely well adapted to generate some key properties.

In order to understand immediately where is destruction and where is creation in Equation (7), let us use our Example 1, listed above, without capital depreciation (with  $\delta = 0$ ). The corresponding DDE is:

$$K'(t) = i(t) - i(t - T(t))(1 - T'(t)). \quad (8)$$

The stock of capital increases with new investments,  $i(t)$  (*creation*), but decreases with scrapped machines,  $i(t - T(t))$ , and with reductions in

the machines' lifetime,  $1 - T'(t)$  (*destruction*). As such, the DDE makes clear the economic determinants of investment dynamics and allows for a precise and comprehensive interpretation. Notice that in the general case (1), differentiation typically yields non-autonomous DDEs reflecting the effects of time-varying vintage-specific weights  $\psi(v, t)$  on the magnitudes of creation and destruction. The whole DDE can be thus highly sophisticated as we will show in our demographic application in the next section.

At the minute, it is worth pointing out that the obtained creation-destruction motions are likely to generate much richer dynamics than the models which omit the vintage-specificity of the populations under consideration. Let us come back to our Example 1, without capital depreciation ( $\delta = 0$ ). If we assume no vintage-specificity, then there is no rationale to scrap machines whatever their age. There is no destruction, and the law of motion of aggregate capital is simply:

$$K'(t) = i(t). \quad (9)$$

With vintage-specificity inducing destruction, capital dynamics are governed by the creation-destruction motion. Equation (8), which can be rewritten as:

$$K'(t) = i(t) - \delta(t) K(t), \quad (10)$$

with:

$$\delta(t) = \frac{i(t - T(t))(1 - T'(t))}{K(t)}. \quad (11)$$

Destruction can be represented as a depreciation scheme. In this case, the depreciation is purely economic (since physical depreciation is assumed zero). Notice that in contrast to the typical law of motions of stock variables, the depreciation rate is not only time-dependent, it is *state-dependent*. It is clear at glance that when vintage-specificity is accounted for, the induced dynamics are likely to be much richer. Without vintage-specificity, the aggregate capital stock is hardly non-monotonic. With vintage-specificity, the capital stock may successively increase and decrease depending on the evolution over time of the magnitude of destruction compared to creation. Non-monotonicity is possible specially because the induced creation-destruction motions entail non-exponential and state dependent depreciation schemes, a fact first observed by Benhabib and Rustichini (1991). Broadly speaking, we claim that such an approach, fully accounting for vintage-specificity, is likely to give rise to the non-monotonic relationships observed either in time series or cross-section analysis. And to make the point, we show in Section 3 that it effectively allows to replicate the typical non-monotonic relationships recently pointed out in the literature of demography and growth (see, for example, Kelly and

Schmidt, 1995). In the next section, we summarize the results of an early application of this approach.

### 2.3 Early Applications: Non-monotonic Investment Paths and Persistent Fluctuations

One of the main failures of the mainstream macroeconomic theory, namely the neoclassical theory, is its incapacity to generate simple models that give rise to persistent output fluctuations, which is one of the most robust stylized facts in the literature (see Cogley and Nason, 1995, for example). The canonical real business cycles models have no persistent propagation mechanisms of the shocks affecting the economies under study. The reason behind that is well-defined: The neoclassical growth model, on which real business cycles models are based, typically displays monotonic optimal paths. And so is the neoclassical growth model because it is built on the assumption that capital is homogenous, namely that there is no vintage specificity, except through a physical capital depreciation at a constant rate. In particular, it is assumed that technological innovations affect identically all the vintages, and therefore there is no reason to scrap machines. The law of motion of capital becomes simply:

$$K'(t) = i(t) - \delta K(t), \quad (12)$$

with  $\delta$  the constant rate of physical depreciation. This makes a big difference with the model accounting for vintage-specificity seen just above where the depreciation rate is not only time-dependent but also state-dependent.

Following Benhabib and Rustichini (1991), we have explored, mainly in optimal control set-ups, whether the modelling of the vintage structure using the DDE approach explained in the previous subsection does effectively display the persistent fluctuations that are missing in the basic neoclassical theory. We did it under stationary environments, that is, without external shocks, as in Boucekkine, Germain and Licandro (1997) or Boucekkine et al. (1998), and under non-stationary environments in Boucekkine, del R o and Licandro (1999). Before giving some details on the outcomes of the latter research, it is worth pointing out that the idea that vintage-specificity may generate oscillatory dynamics traces back to Solow et al. (1996): These authors argue that if the vintages are differentiated by their productivity, or equivalently if technological progress is embodied in capital goods, then the resulting obsolescence of the oldest vintages and their replacement over time may generate persistent fluctuations in investment, and thus in output. In particular, if replacement takes place at regular time intervals, the fluctuations may follow the echo



principle. Episode of high (Resp. low) investment reproduce themselves in the future at their replacement time, giving rise to the so called “replacement echoes.” Solow et al. (1996) did not find such an occurrence in the vintage model they studied, a feature confirmed by Boucekkine, Licandro and Paul (1997) using numerical simulations. However, Solow et al. (1966) did not use an optimal control set-up, and we show hereafter that this makes the difference.

### ***The Generic Optimal Control Vintage Capital Problem***

Consider the following optimization problem:

$$\max \int_0^{\infty} u[c(t)]e^{-\rho t} dt \quad (13)$$

subject to

$$y(t) = \int_{t-T(t)}^t i(z) dz, \quad (14)$$

$$\int_{t-T(t)}^t i(z) e^{-\gamma z} dz = 1, \quad (15)$$

$$y(t) = c(t) + i(t), \quad (16)$$

$$0 \leq i(t) \leq y(t) \quad (17)$$

given the initial conditions  $i(t) = i(t)$  for all  $t < 0$ . The objective of the optimizer is to maximize the discounted sum of a utility stream derived from consumption,  $c(t)$ . As usual in economics, the utility function  $u[c(t)]$  is increasing and concave,  $\rho$  being the time discounting rate. The first constraint is the production function. As in Solow et al. (1996), the production function is linear, of the Leontief type.<sup>1</sup> Each machine of vintage  $z$  is assumed to produce one unit of output. However, the labor required to operate a machine decreases exponentially at a rate  $\gamma$  with the vintage index: technological progress is labor-saving and it is embodied in the successive generations of capital. Therefore, the integral term appearing on

<sup>1</sup>Boucekkine and Pommeret (2004) use a strictly concave production function within a firm problem but obtain the same type of oscillatory investment and output paths. Thus, our results are certainly not due to the Leontief specification. However, replacement echoes need not be perpetual. For example, imperfect competition, by adding concavity to the profit functions, as in Feichtinger et al. (2003), may dampen echo fluctuations.

the lefthand side of the second constraint is precisely labor demand to run all the machines in the workplace, from the oldest, with age  $T(t)$ , to the newest. The second constraint is simply the equilibrium condition in the labor market, where labor supply has been normalized to 1.<sup>2</sup> It is also identical to the labor market equilibrium condition specified in Solow et al. (1966). The third constraint is the resource constraint, and the last constraints are typical feasibility conditions, investment is positive and should not exceed output.

One can use the resource constraint to eliminate  $c(t)$  from the optimization problem and end up with a problem with two controls,  $i(t)$  and  $T(t)$ . It is then relatively easy, using a Lagrangian technique due to Malcomson (1975), to find the following first order condition for an interior maximum:

$$u'(y(t) - i(t)) = \int_t^{t+J(t)} \left(1 - e^{\gamma(z-t-T(z))}\right) u'(y(z) - i(z)) e^{-\rho(z-t)} dz \quad (18)$$

$$J(t) = T(t + J(t)). \quad (19)$$

By construction,  $J(t)$  is the (expected) lifetime of the machines of vintage  $t$ . Equation (18) is the optimal interior investment rule, equalizing the marginal cost of investing in terms of utility at  $t$  (the lefthand side of Equation (18) since one unit more of investment is one unit less for consumption) and the marginal profit from investing in terms of utility, which is the inter-temporal sum of net benefits from investment over the expected lifetime of the purchased capital good. Notice that inter-temporal optimization induces *advanced* integral equations, a feature which makes the control of delayed systems particularly hard. Differentiating Equation (18) together with Equations (14) and (15) yields a differential system with lagged and advanced terms at the same time, and both the lag ( $T(t)$ ) and the lead ( $J(t)$ ) are endogenous.

In the general case, there is no way to solve the problem analytically, and we resort to numerical algorithms. We will come back to our algorithmic solution in the technical section 4. Here, we just give the main computational findings. First, we find that whatever the initial conditions, the optimal paths converge to a steady state state solution where  $c(t)$ ,  $i(t)$  and  $y(t)$  grow at the same rate  $\gamma$ , and where  $T(t) = J(t) = \bar{T}$ , with  $\bar{T}$  a positive constant. More important, and contrary to Solow et al. (1966), the con-

<sup>2</sup>In Boucekkine, del Río and Licandro (1999), we introduce unemployment as an endogenous variable. We get, however, the same type of oscillatory behavior. Therefore, normalizing labor supply to 1 is innocuous from the qualitative point of view. Nonetheless, the considered labor market frictions dampen the echo fluctuations.

vergence to the steady state is oscillatory, featuring the replacement echoes referred to by these authors. That is optimal control makes the difference. In order to understand why, let us solve a special case of the general model above; precisely, we consider linear utility.

### ***The Linear Ramsey Vintage Capital Model***

Linear utility, i.e.,  $u[c(t)] = c(t)$ , yields interior and corner solution as well. In particular, one can show (see Boucekkine, Germain and Licandro, 1997, page 338) that:

$$i(t) \begin{cases} = 0 & \text{if } \Phi(t) < 1 \\ = y(t) & \text{if } \Phi(t) > 1 \\ \in [0, y(t)] & \text{if } \Phi(t) = 1 \end{cases} \quad (20)$$

where

$$\Phi(t) \equiv \int_t^{t+J(t)} (1 - e^{-\gamma(t-z+T(z))}) e^{-\rho(z-t)} dz. \quad (21)$$

and

$$J(t) = T(t + J(t)). \quad (22)$$

Notice that  $\Phi(t)$  corresponds to the righthand side of the general optimal investment rule (18) when utility is linear: It corresponds thus to the weighted stream of profits from investing. Also, in the linear utility case, the marginal cost from investing is 1. Hence, the equations just above simply formalize the intuitive results that optimal investment should be nil (Resp. equal to total resources,  $y(t)$ ) when the benefit from investment is lower (Resp. larger) than the cost. In case of equality between the cost and the benefit, we get the interior solution.

Boucekkine, Germain and Licandro (1997) establish the following results:

- i) *The Terborgh-Smith result.* If the economy is sometime at the interior solution, then it stays on, and both  $T(t)$  and  $J(t)$  are equal to the same constant,  $T^o$ .
- ii) *Finite adjustment Period.* Starting with any initial investment profile, or alternatively with  $T(0)$  not necessarily equal to  $T^o$ , the economy converges at a finite distance to the interior solution.
- iii) *Replacement echoes.* The dynamics of investment and production are characterized by *everlasting* replacement echoes.

It is quite easy to understand why optimal control makes the difference. Property i) implies that the optimizer prefers a constant scrapping time,

which is not surprising since the environment is stationary in this model. But if the optimal scrapping time is constant, then investment and output should fluctuate forever. Property ii) implies that the constant scrapping solution is always feasible from a finite date, whatever the initial investment profile. Henceforth, by simple differentiation of Equation (15), and denoting by  $t_0$  the date at which the interior regime is reached, we have for  $t \geq t_0$ :  $i(t) = i(t - T^o)e^{\gamma T^o}$ , or more eloquently:

$$\hat{i}(t) = \hat{i}(t - T^o), \quad (23)$$

with  $\hat{i}(t) = i(t)e^{-\gamma t}$ . That is, under linear utility, detrended investment is purely periodic of period  $T^o$ , the optimal (interior) scrapping time. Given that the solution is periodic from  $t = t_0$ , it depends on the investment path up to  $t_0$ . More important, any oscillatory feature in the investment path on the interval  $[t_0 - T^o, t_0]$  will be reproduced again in the next future. This is an archetypal occurrence of the replacement echoes referred to before. Naturally, when utility is strictly concave, the fluctuations, are not so regular (because optimal scrapping is no longer constant), and they are dampened. Nonetheless, the oscillations are always there in the short/medium run following the frequency of replacement investment.

### 3 NON-MONOTONIC RELATIONSHIPS BETWEEN DEMOGRAPHIC VARIABLES AND ECONOMIC GROWTH

The relationship between demographic trends and economics is an area of research that is now expanding quickly. The importance of the economic growth process in fostering improvements in longevity has been stressed by the literature, but the feedback effect of past demographic trends on growth, and in particular on the take-off of the Western World, remains largely unexplored. One likely channel through which demographics affect growth is the size and quality of the work force. In this view, generations of workers can be understood as being vintages of human capital, and studied with the same tools as vintages of physical capital.

An interesting point stressed by the empirical literature is that the relation between demographic variables, such as mortality, fertility and cohort sizes, is anything but linear. Kelley and Schmidt (1995) highlight the ambiguous effect of crude death rates. Indeed, growth is slowed by the deaths of the workers but can be enhanced by the deaths of dependents. They provide several elements showing the importance of age-specific mortality rates. Crenshaw, Ameen and Christenson (1997) regress economic growth rates on age-specific population growth rates and conclude that “economies lie fallow during baby booms, but grow rapidly as boomers age and take up their economic roles in societies.”

These non-linear relationships stress the need to model the vintage structure of population. A key element is that different generations have different learning experiences and that the aggregate stock of human capital is built from the human capital of the different generations. In the basic overlapping generation model in continuous time, a new generation is born at each point in time. The members of this generation face a constant probability of death—i.e., independent of age—and the size of each generation declines deterministically through time. De la Croix and Licandro (1999) analyze in this context the optimal education choice; they show in particular that drops in the probability of death induce longer schooling. They also find that the effect on growth is ambiguous, rising life expectancy being good for growth for an economy starting with a high mortality rate, but can be bad for growth in more advanced societies.

In these models, the demographic structure is highly simplified. The survival law is convex and generations never disappear entirely; this is the perpetual youth property implied by a constant probability of death. A richer model is proposed by Boucekine, de la Croix and Licandro (2002). To stress the specific role of the different cohorts, they assume that agents optimally choose the length of the three following activities: learning, working and being retired. Each individual has thus to decide on the length of time devoted to schooling before starting to work and on the retirement age. This model includes a relatively rich description of demographics and a realistic but still tractable survival law. This realistic survival law and the fact that the probability of death is increasing with age clearly represents an improvement compared to previous papers in the field.

### 3.1 The Model

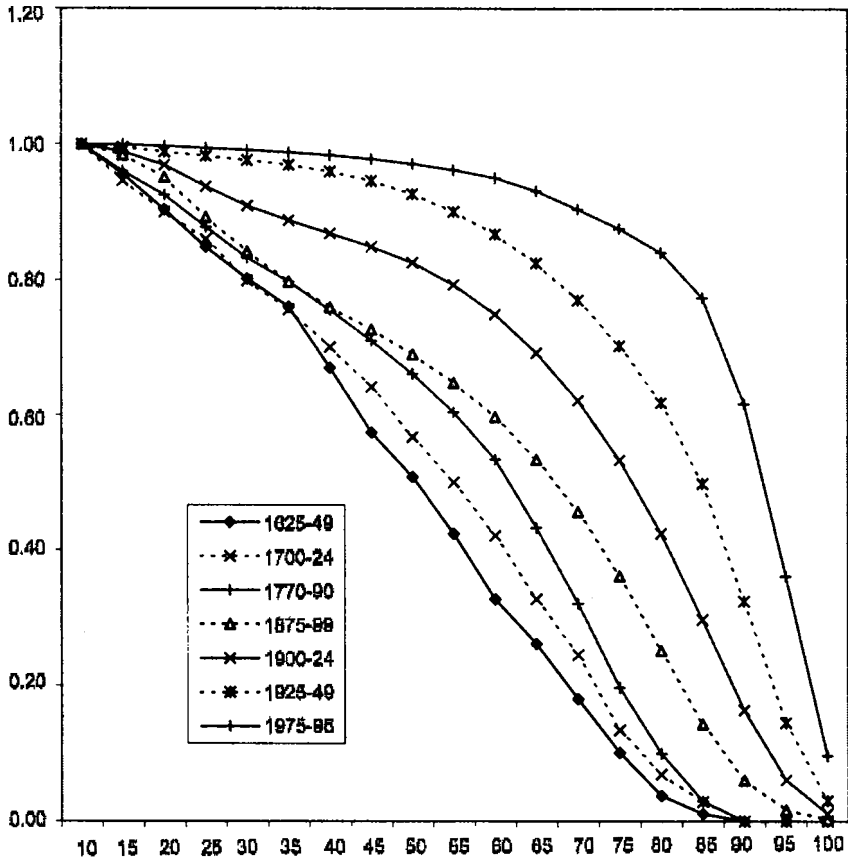
The set of individuals born in  $t$  has a size  $\zeta e^{nt}$  where  $\zeta$  is a scale parameter and  $n$  is the growth rate of population. The probability at birth of surviving at least until age  $\alpha$  is given by:

$$m(\alpha) = \frac{e^{-\beta(\alpha)} - \alpha}{1 - \alpha}. \quad (24)$$

We see here that the survival law depends on two parameters,  $\alpha > 1$ ,  $\beta < 0$ , and is a concave function of age. There is thus an upper bound on longevity obtained by solving  $m(A) = 0$ :  $A = -\log(\alpha)/\beta$ . Life expectancy at birth is given by:

$$\Lambda = \frac{1}{\beta} + \frac{\alpha \log(\alpha)}{(1 - \alpha)\beta} \quad (25)$$

The considered survival law is rich enough to mimic a variety of patterns. Figure 1 puts back to back the survival laws computed for Geneva by



**FIGURE 1** Shifts in the survival law over four centuries.

Perrenoud (1985) over the period 1600–1825 with the French estimates by Vallin and Meslé (2001) over the period 1800–2000. Notice that each curve is normalized so as to start with a value of 1 at age 10, which allow to remove the effect of swings on infant mortality and concentrate on adult mortality. Table 1 presents the estimations of the parameters as well as maximum age and life expectancy (at age 10) for these curves.

The size of the population at time  $t$  is given by

$$\int_{t-A}^t \zeta e^{nz} m(t-z) dz = \zeta e^{nt} \kappa \text{ with } \kappa = \frac{n(1-\alpha) - \alpha\beta(1-\alpha^{n/\beta})}{n(1-\alpha)(n+\beta)} \quad (26)$$

Computing the fertility rate as the ratio of the new cohort to total population we find that it is equal to  $1/\kappa$ . Hence, given the two parameters of the

**TABLE 1** Survival Curve Estimates

Year	$-\beta$	$\alpha$	$10 + A$	$10 + \wedge$
1625–49	0.0091	1.90	70.44	38.97
1700–24	0.0131	2.59	72.56	41.95
1770–90	0.0211	4.78	74.02	46.28
1875–99	0.0195	4.92	81.89	51.39
1900–24	0.0316	15.11	86.00	60.42
1925–49	0.0517	94.83	88.00	69.60
1975–96	0.0858	2394.22	90.69	79.07

survival law  $\{\alpha, \beta\}$ , we can fix  $n$  and deduce the fertility rate  $1/\kappa$ , or alternatively, fix  $\kappa$ , from which we deduce the growth rate of population  $n$ . Notice that  $\partial\kappa/\partial n < 0$ , which reflects the positive relationship between fertility and population growth.

### 3.2 Households' Choices

An individual born at time  $t$  maximizes the following utility function:

$$\int_t^{t+A} c(t, z) m(z-t) e^{\theta(t-z)} dz - \frac{\overline{H}(t)}{\phi} \int_t^{t+P(t)} (z-t) m(z-t) e^{\theta(t-z)} dz, \quad (27)$$

subject to its inter-temporal budget constraint:

$$\int_t^{t+A} c + (t, z) R(t, z) dz = \int_{t+T(t)}^{t+P(t)} h(t) w(z) R(t, z) dz. \quad (28)$$

and to the rule of accumulation of human capital:

$$h(t) = \frac{\mu}{\eta} \overline{H}(t) T(t)^\eta. \quad (29)$$

The choice variables are consumption  $c(t, z)$ , schooling length  $T(t)$ , and retirement age  $P(t)$ . The term  $\overline{H}(t)/\phi \int_t^{t+P(t)} (z-t) m(z-t) dz$ , where  $\overline{H}(t)$  represents average human capital (formally defined in Section 3.4 below), is the disutility of working, which is increasing with age. The parameter  $\phi$  is negatively related to this disutility.  $R(t, z)$  is the contingent price of the consumption good, i.e., the price at time  $t$  for buying one unit of good at time  $z$  conditional on being alive at time  $z$ . The parameter  $\mu$  measures the efficiency in the production of human capital, and the parameter  $\eta$  is the elasticity of income with respect to an additional year of schooling.  $\theta$  represents the time discounting rate.

The first-order conditions to the problem are:

$$m(z - t)e^{\theta(t-z)} - \lambda(t)R(t, z) = 0 \tag{30}$$

$$\overline{H}(t) \left[ \frac{P(t)}{\phi} m(P(t))e^{-\theta P(t)} - T(t)^\eta \lambda(t) R(t, t + P(t)) \frac{\mu}{\eta} w(t + P(t)) \right] - v(t) = 0 \tag{31}$$

$$v(t) \geq 0, \quad P(t) \leq A, \quad v(t)(P(t) - A) = 0 \tag{32}$$

$$\eta T(t)^{\eta-1} \int_{t+T(t)}^{t+P(t)} w(z)R(t, z)dz - T(t)^\eta R(t, t + T(t))w(t + T(t)) = 0. \tag{33}$$

where  $\lambda(t)$  is the Lagrange multiplier of the inter-temporal budget constraint and  $v(t)$  is the Kuhn-Tucker multiplier associated to the inequality constraint  $P(t) \leq A$ .

Let us assume that that the production technology is linear in human capital input:

$$Y(t) = H(t), \tag{34}$$

which implies:  $w(t) = 1, \forall t$ . The real wage par unit of human capital,  $w(t)$  should be equal to 1 to ensure equilibrium in the labor market.

### 3.3 Optimal Schooling and Retirement Decisions

We have studied the properties of this model under various assumptions about parameters. In the simplest case, i.e.,  $\alpha = 0, \eta = 1, \phi = +\infty$  (as in de la Croix and Licandro, 1999), we can solve the model explicitly for optimal schooling. It is given by:

$$T(t) = T = \frac{1}{\beta + \theta}. \tag{35}$$

We clearly see here that improvements in longevity (drop in  $\beta$ ) raise the optimal length of schooling.

Moving toward a richer survival curve with  $\alpha > 1$  and  $\beta < 0$ , allowing for disutility of work and retirement with  $\phi$  finite, and taking  $\theta = 0$  for simplicity, we find in Boucekkine, de la Croix and Licandro (2002) that optimal schooling and retirement depend crucially on the product  $\rho = \mu \phi$ , which can be seen as the ratio of the productivity of schooling to its cost in terms of disutility. The optimal schooling and retirement decisions are given by:

$$P(t) = \min[T(t)\mu \phi, A], \tag{36}$$



$$\int_{t+T(t)}^{t+P(t)} m(z-t)dz - T(t)m(T(t)) = 0. \quad (37)$$

and we have the following proposition:

**Proposition 1:**

- (i) There exists a unique interior  $T^*$  and  $P^* = \rho T^*$  if and only if  $2 < \rho < \rho^*$ .
- (ii) If  $\rho \geq \rho^*$ ,  $T^* = T_{max}(\rho^*)$  and  $P^* = A$ .
- (iii) If  $1 < \rho \leq 2$ ,  $T^* = P^* = 0$ .

In this context we also proved that:

**Proposition 2:** *A rise in life expectancy increases the optimal length of schooling.*

The model was next extended to allow for a non-unit elasticity of income to schooling i.e.,  $\eta < 1$ , in Boucekkine, de la Croix and Licandro (2002).

One implication of the model is that the retirement age will increase with life expectancy. In the data, however, the effective retirement age tended to drop in the OECD countries, while life expectancy rose. This is due to the tax system which incites people to retire early. An alternative specification of the model is to consider  $P(t)$  to be exogenous, say policy determined. The model can then be used to study the effect of changes in the retirement age on schooling and growth.

### 3.4 The Balanced Growth Path

The productive aggregate human capital stock is computed from the capital stock of all generations currently at work:

$$H(t) = \int_{t-P(t)}^{t-\overline{T(t)}} \zeta e^{nz} m(t-z)h(z)dz, \quad (38)$$

where  $t - \overline{T(t)}$  is the last generation that entered the job market at  $t$  and  $t - P(t)$  is the last generation that retired at  $t$ . The average human capital at the root of the externality Equation (29) is obtained by dividing the aggregate human capital by the size of the population given in Equation (26):

$$\overline{H}(t) = \frac{H(t)}{\kappa e^{nt} \zeta}. \quad (39)$$

The dynamics of human capital are thus described by the following delayed integral equation, with delays  $T$  and  $P$ :

$$H(t) = \int_{t-P}^{t-T} m(t-z) \frac{\mu \Gamma^\eta H(z)}{\eta \kappa} dz \quad (40)$$

As far as the long-run is concerned, we have three main results. First, the previous integral equation admits constant growth solutions, the growth rate depending on the demographic parameters and on the productivity of schooling.

Second, the long-run relationship between life expectancy and economic growth is non-monotonic. More precisely, we show in Boucekkine, de la Croix and Licandro (2002) the following Propositions (for  $\theta = 0$  and  $\eta = 1$ ):

**Proposition 3:** *A rise in life expectancy through  $\beta$  at given population growth has a positive effect on economic growth for low levels of life expectancy and a negative effect on economic growth for high levels of life expectancy.*

Third, the relationship between population growth and economic growth is non-monotonic:

**Proposition 4:** *Assume that  $0 < T < P \leq A$ . There exists a population growth rate finite value  $n^*$  such that the long run per capita growth rate of the economy reaches its (interior) maximum at  $n^*$ .*

The model can be used to numerically assess the impact of longer life on schooling and growth. Let us consider the two following parameter sets:

1.  $\eta = .6$ ;  $\theta = .05$ ;  $\mu = .23$ ;  $\phi = 45$ ;  $n = 0.005$ .
2.  $\eta = 1$ ;  $\theta = .0999$ ;  $\mu = .138$ ;  $\phi = 75$ ;  $n = 0.005$ .

The first set assumes an elasticity  $\eta$  in conformity with the econometric studies on schooling and wages. The second set assumes a higher  $\eta$ . In this case, we have adjusted  $\theta$  to keep schooling the same in the two sets for the longevity parameters in 1625–49. The parameters  $\mu$  and  $\phi$  have also been adjusted to keep the expressions  $\mu/\eta$  and  $\eta\phi$  the same in both sets. Table 2 gives the estimated length of schooling and the estimated growth rates for the two parameter sets and for the different parameters of the survival curve provided in Table 1. The age of entry into the labor market is  $10 + T(t)$ . The growth rate  $\gamma$  should be interpreted as the steady state growth of output exclusively due to human capital (there is neither physical capital nor technological progress in the model).

In case 1, the rise of longevity accounts for a lengthening of schooling of about  $2\frac{3}{4}$  years and a rise of long-term growth of 0.5%. In case 2, the effect on schooling is smaller but the effect on growth is proportionally larger, thanks to a larger elasticity  $\eta$ . Note that, in the second case, the

**TABLE 2** Schooling Length and Growth

Parameters	Set 1		Set 1	
	T	$\gamma$	T	$\gamma$
1625–49	4.76	0.91	4.76	0.24
1700–24	5.09	1.03	4.93	0.33
1970–90	5.56	1.19	5.18	0.48
1875–99	5.82	1.14	5.28	0.40
1900–24	6.49	1.27	5.59	0.48
1925–49	7.09	1.36	5.83	0.53
1975–96	7.50	1.36	5.95	0.44

improvements in longevity between 1925–49 and 1975–96 have led to a decrease in the long-term growth rate, illustrating Proposition 3.

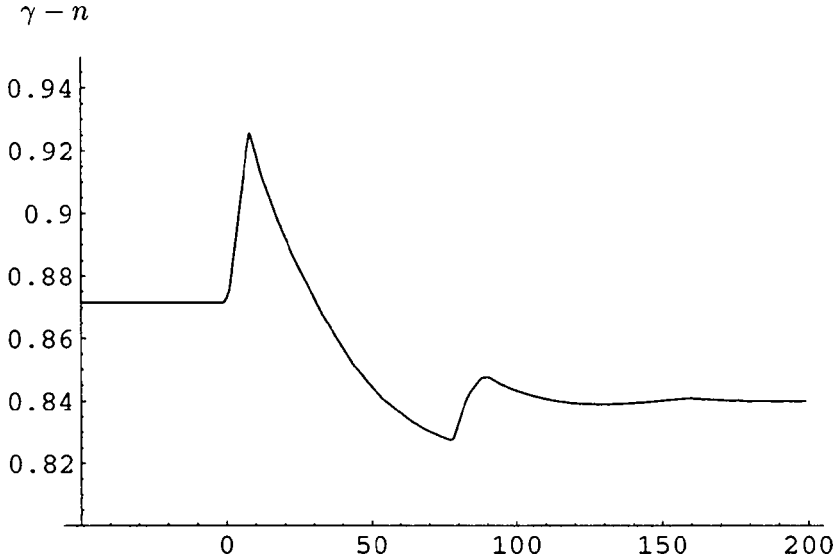
### 3.5 Transitory Dynamics

The transition to the constant growth solutions studied above follows the second-order DDE on  $\hat{H}(t) = H(t)e^{-\gamma t}$ , detrended human capital, where  $\gamma$  is the stationary growth rate:

$$\begin{aligned}
 \hat{H}''(t) &= -\gamma(\beta + \gamma)\hat{H}(t) - (\beta + 2\gamma)\hat{H}'(t) \\
 &+ \frac{\mu\Gamma^\eta}{\eta(1 - \alpha)\kappa} \left[ (\gamma e^{-(\beta+\gamma)T} - \alpha(\beta + \gamma)e^{-\gamma T})\hat{H}(t - T) \right. \\
 &\quad \left. - (\gamma e^{-(\beta+\gamma)P} - \alpha(\beta + \gamma)e^{-\gamma P})\hat{H}(t - P) \right] \\
 &+ \frac{\mu\Gamma^\eta}{\eta(1 - \alpha)\kappa} \left[ (e^{-(\beta+\gamma)T} - \alpha e^{-\gamma T})\hat{H}'(t - T) \right. \\
 &\quad \left. - (e^{-(\beta+\gamma)P} - \alpha e^{-\gamma P})\hat{H}'(t - P) \right]. \quad (41)
 \end{aligned}$$

An example of the transitory oscillatory dynamics is provided in Figure 2. Starting along a balanced growth path, we assume that there is a permanent unexpected change in fertility at  $t = 0$  in an economy with parameter set 1 and longevity parameters for 1975–96. The size of new generations after time zero is  $\zeta$  instead of  $\zeta e^{0.005t}$  for  $t < 0$ . The growth rate of total population changes thus slowly from 0.5% to 0%.

Considering the transition from a balanced growth path to the other, we observe that the change in fertility is first followed by an increase in  $\gamma - \eta$ . Per capita growth rises during 7 years (the schooling length). During this period, the activity rate increases systematically as the weight of students decreases, which exerts a positive effect on growth through the externality. After this period the generations born after  $t = 0$ , which are smaller, start



**FIGURE 2** Dynamic simulation of a fall in fertility.

entering the labor market. This has a negative effect on the externality and dampens growth. After  $t = P$ , the old generations born before 0 are progressively substituted by smaller cohorts in the retired population, which has a positive effect. We then observe small replacement echoes which are typical of models with delays.

## 4 MATHEMATICAL ASPECTS

In this section, we briefly review the technical tools that are relevant in our approach. We also mention some important unsettled issues.

### 4.1 Stability Analysis

There exists an abundant stability literature in the linear case with *constant* delays (Bellman and Cooke, 1963, Hale, 1977, and more recently, Kolmanovski and Myshkis, 1998). General theorems are available for scalar DDEs with a single delay: For example, Hayes theorem (see Theorem 13.8 in Bellman and Cooke, 1963) gives a set of two necessary and sufficient conditions in terms of the coefficients of the DDE (and thus independent of the value of the delay) ensuring that the 0-equilibrium is asymptotically stable (in the sense of Lyapunov). Unfortunately, things become more difficult (and sometimes impossible) when we depart from this simple

class. For example, even in the scalar case, the presence of an additional delay makes it impossible to state general theorems *à la* Hayes, and the values taken by the two delays enter the stability conditions. In any case, as far as the delays or time advances are constant, there is a clear methodology to study stability. We argue hereafter that while this methodology works very well in the absence of optimal control, *it needs to be complemented in the set-ups involving intertemporal optimization as in our approach.*

### ***The Typical Treatment: The Scalar One-delay Case***

To illustrate briefly the typical treatment in the absence of dynamic optimization, consider the Leontief technology (14), and assume investment in new vintages is a constant fraction of output, say  $0 < \alpha < 1$ . Also assume that the lifetime of machines is constant, equal to  $T$ : this is called a one-hoss shay vintage mode, Time differentiation then yields:

$$i'(t) = \alpha(i(t)) - i(t - T)). \quad (42)$$

As for ordinary differential equations, the stability analysis starts with the computation of the roots of the characteristic function  $h(z)$ . Putting  $i(t) = e^{zt}$  in the DDE just above, one gets the characteristic equation:

$$z - \alpha(1 - e^{-zT}) = 0.$$

Here comes the first difficulty of the problem: This is no longer a polynomial problem. Because of the term  $e^{-zt}$ , the lefthand side of the characteristic equation is a transcendental function, and the equation has an infinite number of roots in the set of complex numbers. Precisely, we know that all the roots lie in half a plan with an infinity of roots having negative real parts (see for example, Bellman and Cooke, 1963, chapter 12, for a complete picture). Notice that one could have computed the characteristic equation directly on the structural integral equation, simply by putting  $i(t) = e^{zt}$  in this equation. Observe that the sets of roots of the DDE Vs the integral equation are identical except the root  $z=0$ , which is typically added by time differentiation. However, we find it easier and more comfortable to work on DDEs because of the markedly superior economic insight that can be gained from them as detailed in Section 2, and because of the mostly familiar mathematical concepts and criteria used in comparison with the traditional ordinary differential equation state modelling. This will be clear in a moment.

The infinite number of roots implies two specific problems that are nowadays quite well solved. The first one is purely computational: how to deal with the infinity of roots in practice? Given that only a finite number of roots with a positive real part can occur, a natural numerical solution to the problem is to concentrate on these roots that are responsible for instability.

Some algorithms have been built up to compute these specific roots. For example, Engelborghs and Roose (1999) propose an algorithm, which estimates the subset of rightmost roots of a DDE.

The second problem is more fundamental and has to do with the stability theory itself. Indeed, by analogy with the theory of ordinary differential equations, one would think of the solutions of the DDEs as potentially infinite expansions:

$$i(t) = \sum_r p_r(t) e^{z_r t}, \quad (43)$$

where  $z_r$  is any sequence of roots of  $h$ , and where  $p_r(t)$  is a polynomial of degree less than the multiplicity of  $z_r$ , typically computed from the initial conditions. However, the problem is much harder here because the series are infinite and their convergence is not granted. Fortunately, we have the sufficient material to settle this potential problem quite comfortably (see Theorem 3.4, page 55, and Theorem 4.2, page 109, in Bellman and Cooke, 1963). Not only the series converge under fairly general conditions, but we can extend the familiar Lyapunov theorem for stability and asymptotic stability for linear DDEs. For example, if the roots of a linear homogenous DDE have strictly negative real parts, then 0 is asymptotically stable. Even more familiar existence and Hartman-Grobman theorems hold in the nonlinear scalar case.

### ***Why Optimal Control “Hurts”?***

As pointed out in the sub-section 2.3, inter-temporal optimization yields advanced terms, so that we end up with a mixed-delay system, with both lagged and advanced variables. Even if we abstract away from the state-dependence of the delays and leads, the problem becomes infinitely harder than the pure delay case, or by symmetry, the pure advanced case (i.e., with only time advanced terms). There is a first extreme difficulty in handling mixed-delay systems: As noticed by Rustichini (1988) on a scalar linear mixed-delay equation, such equations may not admit a solution for any initial conditions. For non-zero measure sets of initial conditions, the linear equations considered by Rustichini have no solution.

There is an even more dramatic difficulty: there is no clear stability concept in the mixed-delay case. Indeed, while the delayed part of the system generates an infinite number of roots with negative real parts, its advanced part gives rise symmetrically to an infinite number of roots with positive real part. Which stability criteria to enforce in such a case? What could be a saddlepoint path in this context? This issue is not treated specifically in the traditional literature of delay differential equations. We next propose two ways to tackle it. First, we highlight the role of transversality conditions (whenever necessary) to clarify the stability issue. Then, we present briefly a numerical

algorithm designed to compute the solution paths of the mixed-delay state-dependent system presented in Section 2.<sup>3</sup>

## 4.2 Optimal Control of DDE Systems

As mentioned in the introduction section, the early economic applications of differential-difference equations did not use optimization. The most serious and comprehensive work mixing optimal control of differential-difference equations in economic and management concerns the optimal control of DDEs with infinite (or continuous) delays, which typically arise in optimal dynamic advertising (Hartl, 1984, and Hartl and Sethi, 1984). In the same line of research, one can also quote the remarkable work of Carlson (1990), who focusses on optimal overtaking paths applied to economic growth theory.

In this kind of framework, only the state variables are delayed. Unfortunately, vintage modelling in economics involves finite and state-dependent delays, and more important, generally gives rise to delayed controls. Therefore, the most serious and comprehensive optimal control literature of delayed economic and management systems is not of much help.

In order to make briefly the point, and offer an alternative view of the problem, let us come back to our optimal control vintage capital growth model of Section 2, and let us assume that the machines have a productivity  $A$  (instead of 1 in Section 2) and a finite and constant lifetime,  $T$ . Denoting by  $k(t) = \int_{t-T}^t i(z)dz$ , the problem can be rewritten as follows:

$$\max \int_0^\infty \frac{[Ak(t) - i(t)]^{1-\sigma}}{1-\sigma} e^{-\rho t} dt \quad (\text{P})$$

subject to

$$k'(t) = i(t) - i(t - T), \quad (44)$$

$$0 \leq i(t) \leq Ak(t), \quad (45)$$

given  $i(t) = i_0(t) \geq 0$  for all  $t \in [-T, 0]$ , and

$$k(0) = \int_{-T}^0 i_0(z)dz. \quad (46)$$

<sup>3</sup>We have not explored so far the semi-group theories recently developed to analyze the class of infinite dimensional ordinary differential equations, to which DDEs belong. See for example Bensoussan et al., 1992. We thank an anonymous referee of this journal for bringing our attention to this literature.

As mentioned above, the problem turns out to be an optimal control problem with delayed controls, and not with delayed state variables as in the core of the related literature. Nonetheless, some literature (Kolmanovski and Myshkis, 1998, chapter 14, or even Kamien and Schwartz, 1991, Section 19) exists on problems like ours but only in finite horizon. Boucekkine et al. (2003) invoke Michel's argument (1982) to identify the transversality conditions that arise as *necessary* optimality conditions in infinite horizon, and use these conditions to conclude for the stability of the optimal paths, therefore surmounting the conceptual problem around the stability of mixed-delays systems mentioned in sub-section 4.1.

Indeed, the first order conditions for an interior solution are:

$$[Ak(t) - i(t)]^{-\sigma} e^{-\rho t} = \lambda(t) - \lambda(t + T) \quad (47)$$

$$A[Ak(t) - i(t)]^{-\sigma} e^{-\rho t} = -\lambda'(t) \quad (48)$$

and the transversality conditions

$$\lim_{t \rightarrow \infty} \lambda(t) \geq 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \lambda(t) \kappa(t) = 0 \quad (49)$$

The transversality conditions are *necessary* because following Michel (1982), the objective function is positive and because the set of admissible speeds of the optimal state variable, namely  $k'(t) = i(t) - i(t - T)$ , for all possible controls  $i(t)$  contains a neighborhood of 0 for  $t$  large enough. Indeed,  $k'(t) \in [-Ak^*(t - T), Ak^*(t)]$  and  $k^*(t)$  does not tend to 0 (Boucekkine et al., 2003, Proposition 8). A sufficiency proof *à la* Mangasarian is also easy to obtain (Boucekkine et al., 2003, Proposition 9).

The linearity of the production function allows to isolate a purely advanced differential (ADE) equation in the co-state  $\lambda(t)$ :

$$\lambda'(t) = A(\lambda(t + T) - \lambda(t)), \quad (50)$$

Hence, in this special case, we are able to handle the mixed-delay system sequentially: first the ADE part (the dynamics of the co-state), then the DDE part (the dynamics of capital accumulation). As mentioned in the previous sub-section, the first one generates an infinity of roots with positive real part, and the second an infinity of roots with negative real part. And here comes the fundamental role of the transversality conditions as traditionally in economic dynamics, namely the selection of stable paths. As shown in Boucekkine et al. (2003)(Lemma 14), the transversality conditions (49) do eliminate the destabilizing roots coming from the ADE (50), inducing (asymptotically) stable paths.

Overall, the transversality conditions should play the same role in the optimal control of differential-difference systems as in the control of ordinary differential equations. The recent literature of optimal overtaking paths, mentioned above, does not take this approach. In the PDE approach



advocated by Feichtinger et al. (forthcoming), the transversality conditions are even not necessary for the class of models considered by these authors.<sup>4</sup> In our approach, such conditions seem to have a natural role, though one should acknowledge that their optimality (specially as *necessary* conditions) is not always granted. In the special case examined in Boucekkine et al. (2003), our story works because the linearity of our state equation allows to use Michel's argument relatively easily. In less special cases, this may be a daunting task.

### 4.3 Algorithms

In the presence of state-dependent delays, the analytical analysis of the dynamics becomes almost impossible except in some special cases (like the linear vintage models explored in the second and third sections of this paper). Therefore, we must resort to numerical algorithms. State-dependent delay differential (or integro-differential) equations can be solved safely using some refinements of the method of steps (see for example, Baker and Paul, 1993, for one of these extensions, and Boucekkine, Licandro and Paul, 1997, for an easy exposition with economic applications). Unfortunately, there are few algorithms designed to solve mixed-delay systems, arising from the optimal control of DDE systems, most probably because mixed-delay equations do not occur so frequently in the natural world, which motivates most computational mathematics. One of the exceptions is Chi, Bell and Hassard (1986), who use finite elements methods to solve a nonlinear advance-delay differential equation from nerve conduction theory. However, the delay and time advance are constant in this paper.<sup>5</sup>

A feasible technique to deal with the nonlinear optimal control-based vintage models seen in Section 2 is to solve *directly* the optimization problem, that is without using the first-order necessary optimality conditions, which cause the state-dependent delays and advances to occur simultaneously. An example of this technique is Boucekkine et al. (2001), and an additional (early) application of it is in Boucekkine et al. (1998). Recall the generic problem:

$$\max \int_0^{\infty} u[y(t) - i(t)]e^{-\rho t} dt \quad (51)$$

subject to

<sup>4</sup>Of course, this should not be taken as an indication that transversality conditions are systematically not necessary in the control of PDE systems.

<sup>5</sup>See also Collard, Licandro and Puch (2004).

$$y(t) = \int_{t-T(t)}^t i(z) dz, \tag{52}$$

$$\int_{t-T(t)}^t i(z) e^{-\gamma z} dz = 1, \tag{53}$$

with  $0 \leq i(t) \leq y(t)$ , and given  $i(t) = i(t)$  for all  $t < 0$ . One can handle it numerically as follows:

**Discretisation:** Replace the unknown functions  $i$  and  $y$  by *piecewise constant functions* on the intervals  $(0, \Delta)$ ,  $(\Delta, 2\Delta)$ ,  $\dots$ . Let  $i_0, i_1, \dots; y_0, y_1, \dots$  be the unknown values. Discretise the objective function and the integral constraints (14)-(15) as well. For example, the integral objectives function may be performed as

$$\sum_{k=0}^N u(y_k - u_k) \frac{e^{-k\rho\Delta} - e^{-(k+1)\rho\Delta}}{\rho}. \tag{54}$$

**Maximization by iteration:** Maximize the integral by iteration, starting with an initial investment vector  $[i_0 i_1 i_N]'$ , the base of the relaxation. Then:

- i) **step 0** Maximize the discretized integral objective function with respect to  $i_0$  keeping unchanged all the subsequent investment ordinates with respect to the base. Update  $i_0$  with the resulting maximizand.
- ii) **step k** for  $k = 1, \dots, N$ , maximize the discretized objective function with respect to  $i_k$  keeping unchanged the posterior investment ordinates, if any, with respect to the base, with the anterior investment ordinates  $i_l, 0 \leq l \leq k - 1$ , updated thanks to the anterior maximization steps.
- iii) Update  $i_k$  using the resulting maximizand. Update the investment vector. Update the vector  $(T(t_k)), k = 1$  to  $N$ , using the discretized (15) and the updated investment vector. Update the vector  $(y^k), k = 1$  to  $N$  using the discretized (14) and the updated investment and scrapping time vectors.

**Relaxation iteration:** Redo i) to iii) until convergence of the investment vector.

The maximization by iteration device included in this algorithm corresponds to the cyclic coordinate descent optimization algorithm described in Luenberger (1965), pp.158–1961. The algorithm is especially useful because of its easy implementation as it does not require any information on the gradient of the objective functions in contrast to most alternative

methods. In any case, it has proved reliable and useful in all the numerical work we have conducted so far.

## 5 CONCLUSION

In this paper, we have presented the DDE approach to modelling vintage structures, with some applications to make clear its contributions and its limits. We have shown how the optimal control ingredient is essential to the approach for a comprehensive treatment of the economic problems under study. We have also shown how this ingredient complicates tremendously the analysis. Nonetheless, this paper shows clearly that the approach allows for a relatively complete and rigorous analytical exploration in some special cases (mainly linear or quasi linear models), and for an easy computational appraisal in the general case.

However, these performances should not hide the fact that the optimal control of DDEs poses some tremendously difficult problems in less generic cases, including existence problems. These issues are on the top of our agenda.

## REFERENCES

- Baker, B. and Paul, C. (1993). Parallel continuous Runge-Kutta methods and vanishing delay differential equations. *Advances in Computational Mathematics* 1, 367–394.
- Barucci, E. and Gozzi, F. (2001). Technology adoption and accumulation in a vintage capital model. *Journal of Economics* 74, 1–38.
- Bellman, R. and Cooke, K. (1963). *Differential-Difference Equations*. Academic press.
- Benhabib, J. and Rustichini, A. (1991). Vintage capital, investment, and growth. *Journal of Economic Theory* 55, 323–339.
- Bensoussan, A., Da Prato, G., Delfour, M., and Mitter, S.K. (1992). *Representation and Control of Infinite Dimensional Systems*. Volume 1 and 2. Systems and Control: Foundations and Applications. Birkhäuser.
- Boucekkine, R., Germain, M., and Licandro, O., (1997). Replacement echoes in the vintage capital growth model. *Journal of Economic Theory* 74, 333–348.
- Boucekkine, R., Licandro, O., and Paul, C. (1997). Differential-difference equations in economics: on the numerical solution of vintage capital growth models. *Journal of Economic Dynamics and Control* 21, 347–362.
- Boucekkine, R., Germain, M., Licandro, O., and Magnus, A. (1998). Creative destruction, investment volatility and the average age of capital. *Journal of Economic Growth* 3, 361–384.
- Boucekkine, R., de Río, F., and Licandro, O. (1999). Endogenous Vs exogenously driven fluctuations in vintage capital models. *Journal of Economic Theory* 88, 161–187.
- Boucekkine, R., Germain, M., Licandro, O. and Magnus, A. (2001). Numerical solution by iterative methods of a class of vintage capital models. *Journal of Economic Dynamics and Control* 25, 655–699.
- Boucekkine, R., de la Croix, D., and Licandro, O. (2002). Vintage human capital, demographic trends and endogenous growth. *Journal of Economic Theory* 104, 340–375.

- Boucekkine, R., de la Croix, D., and Licandro, O. (2003). Early mortality declines at the dawn of modern growth. *Scandinavian Journal of Economics* 105, 401–418.
- Boucekkine, R., Licandro, O., Puch L., and del Río, F. (2003). Vintage capital and the dynamics of the AK model. *Journal of Economic Theory*, forthcoming.
- Boucekkine, R. and Pommeret, A. (2004). Energy-saving technical progress and optimal capital stock: The role of embodiment. *Economic Modelling* 21, 429–444.
- Carlson, D. (1990). The existence of catching-up optimal solutions for a class of infinite-horizon control problems with time delay. *SIAM Journal of Control and Optimization* 28, 402–422.
- Chi, H., Bell, J., and Hassard, B. (1986). Numerical solution of a nonlinear advanced delay-differential equation from nerve conduction theory. *Journal of Mathematical Biology* 24, 583–601.
- Cogley, T. and Nason, J. (1995). Output dynamics in real-business-cycles models. *American Economic Review* 85, 492–512.
- Collard, F., Licandro, O., and Puch, L. (2004). *The short-run dynamics of optimal growth models with delays*. European University Institute, ECO 2004/04.
- Crenshaw, E., Ameen, A., and Christenson, M. (1997). Population dynamics and economic development: age-specific population growth and economic growth in developing countries, 1965 to 1990. *American Sociological Review* 62, 974–984.
- De la Croix, D. and Licandro, O. (1999). Life expectancy and endogenous growth. *Economics Letters* 65, 255–263.
- Engelborghs, K. and Roose, D. (1999). Numerical computation of stability and detection of Hopf bifurcations of steady state solutions of delay differential equations. *Advances in Computational Mathematics* 10, 271–289.
- Feichtinger, G., Hartl, R., Kort, M., and Veliov, V. (forthcoming). Capital accumulation under technological progress and learning: a vintage capital approach. *European Journal of Operation Research*, forthcoming.
- Feichtinger, G., Hartl, R., Kort, M., and Veliov, V. (2003). *Anticipation Effects of Technological Progress on Capital Accumulation: A Vintage Capital Approach*. Mimeo: Institute of Management, University of Vienna.
- Galor, O. and Weil, D. (1999). From Malthusian stagnation to modern growth. *American Economic Review* 89, 150–154.
- Hale, J. (1977). *Theory of functional differential equations*. Springer-Verlag.
- Hartl, R. (1984). Optimal dynamic advertising policy for hereditary processes. *Journal of Optimization Theory and Applications* 43, 51–72.
- Hartl, R. and Sethi, S. (1984). Optimal control of a class of systems with continuous lags: Dynamic programming approach and economic interpretations. *Journal of optimization Theory and Applications* 43, 73–88.
- Kalecki, M. (1935). Macroeconomic theory of business cycles. *Econometrica* 3, 327–344.
- Kamien, M. and Schwartz, N. (1991). *Dynamic Optimization*. North-Holland.
- Kelley, A. and Schmidt, R. (1995). Aggregate population and economic growth correlations: the role of the components of demographic changes. *Demography* 32, 543–555.
- Kolamanovski, V. and Myshkis, A. (1998). *Introduction to the Theory and Applications of Functional Differential Equations*. Kluwer.
- Leontief, W. (1961). Lags and the stability of dynamic systems. *Econometrica* 29, 659–669.
- Luenberger, D. (1965). *Introduction to Linear and Nonlinear Programming*. Addison-Wesley.
- Malcomson, J. (1975). Replacement and the rental value of capital equipment subject to obsolescence. *Journal of Economic Theory* 10, 24–41.
- Michel, P. (1982). On the transversality condition in infinite horizon optimal problems. *Econometrica* 50, 975–986.

- Perrenoud, A. (1985). Mortality decline in its secular setting. In *Pre-industrial population change*. T. Bengtsson, G. Fridlitzius, and R. Ohlsson (eds). Almqvist and Wiksell International 41–69.
- Rustichini, A. (1988). Functional differential equations of mixed type: The linear autonomous case. *Journal of Dynamics and Differential Equations* 1, 121–143.
- Solow, R., Tobin, J., von Weizsacker, C., and Yaari, M. (1966). Neoclassical growth with fixed factor proportions. *Review of Economic Studies* 33, 79–115.
- Vallin, J. and Meslé, F. (2001). *Tables de mortalité françaises pour les XIX<sup>e</sup> et XX<sup>e</sup> siècles et projections pour le XXI<sup>e</sup> siècle*. Paris: INED.

Copyright of Mathematical Population Studies is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.