OPTIMAL LOCALIZED PRODUCTION EXPERIENCE AND SCHOOLING*

BY CHARLES VAN MARREWJIK, CASPER G. DE VRIES, AND CEEE WITHAGEN1

Three factors of economic growth, i.e., physical capital accumulation, schooling and learning by doing, are investigated. The special relationship between the first factor and the other two is modeled through adjustment costs in production experience accumulation whenever the production technique changes. This generates a nontrivial two state variable control problem. Necessary and sufficient conditions for optimality are given. The steady state and the path to it are unique. It is shown that, among others, technique changes deter whereas learning by doing enhances the speed of convergence to the steady state.

I. INTRODUCTION

A recent article, Lucas (1988), discusses three (neo-)classical models of economic growth. The first model emphasizes physical capital accumulation and technological change, the second human capital accumulation through schooling and the third human capital accumulation through learning by doing. The aim of this paper is to investigate these issues simultaneously within one growth model.

This is accomplished by extending the standard neoclassical growth model as in Solow (1956), to one with three inputs: capital, labor and production experience. The latter input represents the endogenous technical change induced by human capital accumulation, compare Nordhaus (1967), which is arrived at by either investing in schooling or through learning by doing. Schooling is now widely recognized as a prime factor of economic growth, see e.g. the discussion in Becker (1981). Before the issue of schooling came into focus, economists realized the importance of on the job training as an alternative channel for accumulating production experience; compare Wright (1936) and Alchian (1963). Here we follow Bardhan (1971), Rosen (1972) and Woodland (1982, Ch. 15) in modeling the learning by doing process. Killingsworth (1982) is the first to study schooling and learning by doing simultaneously in a model of human capital accumulation. (The difference between the two being that one can choose no education, while learning by doing is unavoidable.) In this study we extend Killingsworth's analysis by linking the two human capital growth factors with the third growth factor, i.e. standard physical capital accumulation. This is done as follows. In general the stock of production knowledge is tied to the specific production techniques that are employed. Atkinson and Stiglitz (1969) therefore argue that learning by doing effects are

* Manuscript received September 1989; revised June 1990 and January 1991.
1 We are grateful to P. Braunerhjelm, P. Broer, D. Furth, G. Heal, J. van Geldrop, S. Hu, M. Peeters, M. van de Sande Bakhuysen, T. ten Raa, J. Verbeek, T. Vorst, and the anonymous referees for their stimulating, detailed and valuable comments.
localized. That is, if a firm changes its production technique it incurs a loss in productivity due to a lack of experience with the new production process. Empirical evidence on this matter is presented in the study by Gulledge and Womer (1986). One way of thinking about such localized technical progress is in terms of adjustment costs. See, specifically, Uzawa's (1968) discussion of the Penrose effect and, more recently, the Prescott and Visscher (1980), and Becker (1983) contributions on administrative and managerial adjustment costs. In summary, schooling and learning by doing are considered alternative means for endogenous production experience accumulation, while technique changes retard this process.

The above ideas are formalized and analyzed within an optimal growth model. The various aspects of economic growth imply a two state variable optimal control problem. Despite the complexity of the model due to the simultaneous modeling of the different growth factors, the properties of the model in terms of steady state values and adjustment speed can be characterized. Frequently the integration of different strands of literature into one model leads to ambiguity. With a few minor exceptions this is not the case in our model. For example, we establish existence and uniqueness of the steady state and obtain stability. This allows one to investigate how the steady state consumption is affected by population growth, forgetfulness, learning by doing and schooling expenditures. The importance of schooling is assessed by comparing the economy with another economy which has the same stock of steady state knowledge but no schooling possibilities. Localized learning affects the speed of convergence towards the steady state negatively. In the next two sections the growth problem is formalized and sufficient conditions for optimality are given. Section 4 analyzes the properties of the steady state, Section 5 investigates stability, and Section 6 studies the speed of convergence.

2. PROBLEM STATEMENT

In this section we sketch the model and list the assumptions. Time indices are omitted if there is no danger of confusion. Aggregate production \( Y \) is a linearly homogeneous function of production experience \( P \), capital \( K \) and indispensible labor \( L \):

\[
Y = F[P, K, L].
\]  

(2.1)

Inclusion of a human capital element in the production function alongside physical capital is now a standard procedure, but frequently human capital replaces labor as an independent argument. Here we add human capital for the sake of generality, and because this is common practice in the literature on learning by doing, see e.g. Bardhan (1971) and Woodland (1982). The specification is also in line with the recent empirical evidence on the absence of externalities to the capital input and the importance of a third factor, see e.g. Benhabib and Jovanovic (1991) and Mankiw et al. (forthcoming). But production experience may not be a necessary input. Therefore we assume that while labor is indispensible for production, the other two factors are not indispensible. The motions of capital and labor are standard:

\[
\dot{K} = I - \mu K, \quad K(0) > 0 \quad \text{given},
\]  

(2.2)
\[ \dot{L} = \pi L, \quad L(0) > 0 \quad \text{given}. \]

Here \( I \) denotes gross investments, while \( \mu \) and \( \pi \) are the rates of depreciation and population growth.

The specification of production experience accumulation, however, merits discussion. While there is some literature dealing with production experience accumulation in the aggregate, most literature deals with this process on the individual's level. Therefore we decided to model this process on the per capita level. For any variable \( X \) write \( x := X/L \), to denote per capita levels. The aim of the paper is to integrate two distinct accumulation factors, learning by doing and schooling, previously treated separately in the macro economics literature, as well as to consider the effects of physical capital accumulation on production experience. On the basis of empirical evidence, see e.g. the comprehensive studies by Conlisk (1967, 1970) and Gulledge and Womer (1986), learning by doing is commonly represented through output \( y \). The training of labor and process specific research activities also improve the operation of the production process. Outlays for schooling \( s \) are considered as a main contributor to this type of learning, see e.g. Shell (1966), Ben-Porath (1967), Becker (1975), Heckman (1976) and Moreh (1980). Thus we model schooling as an expenditure allocation problem, i.e. national income is divided over investment in physical capital, educational expenses and consumption.\(^2\) Note that schooling in this setting is a decision variable, whereas the learning by doing effect is endogenous. Both, learning by doing and schooling are enhanced by the amount of existing production experience \( p \), see Ben-Porath (1967), Bardhan (1971), Rosen (1972), Becker (1975) and Lucas (1988). As time passes by some experience is lost due to, for example, retirement. An empirical study by Kipps and Kohen (1984) estimates this loss in the order of 4 to 10 percent per year; see also Benhabib and Jovanovic (1991). The latter two effects are combined into one term \( mp \), where \( m \) is the rate of net forgetfulness. Finally, the effectiveness of schooling and learning by doing is reduced if new production techniques are introduced, see e.g., Gulledge and Womer (1986), and therefore the accumulation process becomes localized as is argued in Atkinson and Stiglitz (1969). We model this effect by means of an adjustment cost \( A, A = A(k) \), i.e. for \( k \neq 0 \) some experience is lost. The technique changes are identified by capital labor ratio changes, as suggested by Atkinson and Stiglitz.

All these factors together influence \( \dot{p} \) in the following way:

\[ \dot{p} = a\dot{y} + bs - mp - A(k), \]

where \( a \) is the coefficient of learning by doing, the schooling coefficient \( b \) captures the marginal efficiency of schooling expenditures and \( y \) is the per capita format of the production function \( F \).\(^3\)

\(^2\) Alternatively the decision to educate is sometimes modeled as a time allocation problem, see e.g., Becker (1975), Killingsworth (1982) or Lucas (1988). Ben-Porath (1967) and Heckman (1976) combine the two approaches on the individual level, while van de Sande Bakhuyzen (1991) provides a combination on the aggregate level.

\(^3\) In van Marrewijk, de Vries and Withagen (1988), a somewhat more general nonlinear specification is considered. Additivity of the different components which generate the production experience is
Gross aggregate output in the economy is allocated to consumption \( C \), investment \( I \) and schooling costs \( S \). Hence, the national income identity reads

\[
Y = C + I + S.
\]

We investigate a plan economy employing a Benthamite utilitarian welfare functional, see e.g., Arrow and Kurz (1970, ch. 1),

\[
\int_{0}^{\infty} e^{-\rho t} LU(C/L) dt,
\]

where \( \rho' (> 0) \) is the constant rate of time preference. The homogeneity assumption enables one to write the entire model in per capita terms. The optimization problem then reads

\[
\max_{c, i, s} \int_{0}^{\infty} e^{-\rho t} U(c) dt
\]

subject to

\[
\dot{p} = \alpha y(p, k) + bs - mp - A(\dot{k}), \quad p(0) > 0 \quad \text{given},
\]

\[
k = i - nk, \quad k(0) > 0 \quad \text{given},
\]

\[
y(p, k) = c + i + s,
\]

\[
c \geq 0,
\]

\[
k \geq 0, \quad p \geq 0,
\]

where \( n = \mu + \pi, \rho = \rho' - \pi \). Without loss of generality, the size of the initial population is set equal to unity. With respect to the functions involved some additional assumptions are made in order to deal with a manageable problem.

A.1. Let \( \Gamma = [0, M_1] \times [0, M_2] \), with \( M_1 \) and \( M_2 \) finite but arbitrarily large (large enough to comprise the steady state introduced in Section 4 below). Then \( y \) is defined on \( \Gamma \) and twice continuously differentiable on \( \text{int}(\Gamma) \), \( y(p, k) \) is strictly concave, \( y_k(p, k) > 0, y_p(p, k) > 0, y_{pk} \geq 0, \forall p > 0 y_k(p, k) \to \infty \) as \( k \to 0 \), \( \forall k > 0 y_p(p, k) \to \infty \) as \( p \to 0 \).

Note that A.1 resembles the familiar Inada conditions. One of its implications is that along an optimal trajectory \( p > 0 \) and \( k > 0 \), so that essentially the problem

---

4 We allow negative schooling to take place. This is analogous to the possibility of negative investment, whereby part of the installed capital stock is consumed. If one wants to dispense with negative schooling an additional constraint, \( S \geq 0 \), has to be taken into account.
contains no pure state constraints, compare Arrow and Kurz (1970, p. 48). Thus restraints (2.12) can be dropped.

A.2. \( A \) is defined on \( \mathbb{R} \), is twice continuously differentiable, and \( A(0) = 0, A'(0) = 0, A'(+\infty) = +\infty, A'' > 0 \).

Adjustment costs are convex and there are no adjustment costs when the production technique does not change.

A.3. \( U \) is defined on \( \mathbb{R}_+ \) and twice continuously differentiable on \( \mathbb{R}_{++} \), \( U' > 0 \), \( U'(c) \rightarrow +\infty \) as \( c \rightarrow 0 \), \( U'' < 0 \).

This assumption implies that along the optimal trajectory consumption is strictly positive, i.e., (2.11) can be omitted.

A.4. \( \rho + m > 0, \rho + n = \rho' + \mu > 0, \rho > 0 \).

This last assumption ensures existence of a steady state.

Mathematically our model is a two state variable optimal control problem. In order to invoke the Pontryagin maximum principle the class of admissible functions is to be restricted. The state variables \( (p, k) \) are continuous and the control variables \( (c, i, s) \) are piecewise continuous.

Before turning to the analysis, it is useful to discuss the relation between the problem in equations (2.7) through (2.10) and the literature. Shell (1966) formulates a two state variable problem similar to ours except for the learning by doing and localization effects; no in depth treatment is provided. Conlisk (1967, 1970) models the learning by doing effect, and is able to reduce the system to a single state variable problem. Heckman (1976) formulates a growth problem taking both educational expenses and schooling time into account, but no physical capital accumulation or learning by doing is considered. Killingsworth (1982) obtains a two state variable problem by modeling learning by doing and schooling as a time allocation problem, i.e. time on the job and time in school, without the localization effect and a simple single input production function; implications are derived in a cursory manner. Some recent empirical evidence for the three factor constant returns to scale specification is given by Mankiw et al. (forthcoming). Given this state of the art, the problem (2.7) through (2.10) is of interest as it considers the three factors of economic growth simultaneously. The analysis below is greatly facilitated by the analytical analogy between our problem and the analysis of heterogeneous capital accumulation in Cass and Shell (1976), see Section 5 below.

3. CONDITIONS FOR OPTIMALITY

Define the current value Hamiltonian \( H(k, p, \lambda, \gamma, c, i, s, \omega, t) = U(c) + \lambda[i - nk] + \gamma[ay(p, k) + bs - mp - A(i - nk)] + \omega[y(p, k) - c - i - s] \).
The following theorem is a straightforward application of Pontryagin's maximum principle (see e.g., Takayama 1974 or Arrow and Kurz 1970). It gives necessary and sufficient conditions for the optimum policy.\(^5\)

**Theorem 3.1.** Suppose that \( [p(t), k(t), c(t), i(t), s(t)] \) is a solution to the problem (2.7) through (2.10). Then there exist continuous \( \gamma(t) > 0 \) and \( \lambda(t) > 0 \), and \( \omega(t) \geq 0 \), possibility discontinuous at points of discontinuity of \([c(t), i(t), s(t)]\), such that (2.8) through (2.10) hold and

\[
\begin{align*}
U'(c(t)) &= \omega(t), \\
\lambda(t) - \gamma(t) A'(i(t) - nk(t)) &= \omega(t), \\
b\gamma(t) - \omega(t) &= 0, \\
\dot{\lambda}(t) &= (\rho + n) \lambda(t) - [\gamma(t)a + \omega(t)] y_k(p(t), k(t)) - \gamma(t) n A'(i(t) - nk(t)), \\
\dot{\gamma}(t) &= (\rho + m) \gamma(t) - [\gamma(t)a + \omega(t)] y_p(p(t), k(t)).
\end{align*}
\]

If there exist \([p(t), k(t), c(t), i(t), s(t), \gamma(t), \lambda(t), \omega(t)]\) such that (2.8) through (2.10) and (3.1) through (3.5) hold, and in addition

\[
\begin{align*}
e^{-\rho t} \lambda(t) k(t) &\rightarrow 0 \quad \text{as} \quad t \rightarrow \infty, \\
e^{-\rho t} \gamma(t) p(t) &\rightarrow 0 \quad \text{as} \quad t \rightarrow \infty,
\end{align*}
\]
then \([p(t), k(t), c(t), i(t), s(t)]\) is optimal. \(\Box\)

The set of necessary conditions consists of the four differential equations (2.8), (2.9), (3.4), (3.5) and four other equations (2.10), (3.1) through (3.3). There are two state variables (\(p\) and \(k\)), two costate variables (\(\lambda\) and \(\gamma\)), three control variables (\(c\), \(i\) and \(s\)) and one multiplier (\(\omega\)). The control variables cannot be chosen independently: choosing any two of these implies the third by the budget constraint. Therefore \(\omega\), the budget constraint multiplier, is always positive, as equation (3.1) shows. The set of sufficient conditions consists of the set of necessary conditions and in addition the two transversality conditions (3.6), and (3.7).

There is a convenient way to reduce the system of necessary conditions.

**Lemma 3.1.** Consider the set of equations (2.10), (3.1) through (3.3). To any \([p(t), k(t), \gamma(t), \lambda(t)]\) there corresponds a unique \([c(t), i(t), s(t)]\).

**Proof.** The proof exploits the monotonicity of the functions as is implied by A.1 through A.3. After substitution of \(\omega(t)\) from (3.1) into (3.2) and (3.3) and by A.3, one sees immediately that \(c(t)\) only depends on \(\gamma(t)\). Given A.2, which implies

\(^5\) For the proof of Theorem 3.1, exploiting the strict concavity of \(U, y, -A\), and the transversality conditions, the reader is referred to van Marrewijk (1988).
that \( A'(\cdot) \) is monotonic, \( i(t) \) follows from (3.2) and (3.3) as a function of \( \gamma(t) \), \( \lambda(t) \) and \( k(t) \). Using these results in (2.10) together with A.1 yields the lemma.

Comparative statics for the instruments with respect to the state and costate variables are given in the Appendix, part 1. Lemma 3.1 enables us to reduce the set of necessary conditions to four differential equations in four variables. This system is analyzed further in the next section.

The costate variables \( \lambda \) and \( \gamma \) can be interpreted as the shadow prices of investment and schooling respectively. Therefore, the Hamiltonian gives the shadow profits in terms of total welfare. The partials \( \partial H/\partial \lambda \) and \( \partial H/\partial \gamma \), given by (2.8) and (2.9), indicate the optimal investment and schooling levels. The partials \( \partial H/\partial k \) and \( \partial H/\partial p \), given in (3.4) and (3.5), represent the marginal value of capital and production experience.

The intuition behind the system (2.8) through (2.10), (3.1) through (3.5) can be developed further by the following economic interpretation. Differentiate equations (3.1) and (3.3) with respect to time and rewrite equation (3.5)

\[
(a + b)y_p - m = \rho + \eta(c) \frac{\ddot{c}}{c},
\]

where \( \eta(c) \) is the elasticity of marginal felicity. The first term on the left-hand side of equation (3.8) is the sum of the marginal productivity of learning by doing and schooling expenditures. The entire left-hand side is the net marginal productivity. For example, a currently low stock of production experience, and hence high marginal productivity, calls for accumulation of \( p \). This requires initially high schooling expenditures and low consumption levels. Over time, however, as production experience accumulates the consumption level can rise, i.e., \( \ddot{c}/c > 0 \).

4. THE STEADY STATE AND ENDOGENOUS HUMAN CAPITAL ACCUMULATION

In this section the steady state is characterized and comparative statics results on the effects of learning by doing and schooling are derived. The case of endogenous human capital accumulation is compared with the case of an exogenously given human capital level. We start with a definition of the steady state.

**Definition 4.1.** A steady state is defined by constant per capita levels of capital and production experience: \( \ddot{p} = \ddot{k} = 0 \).

**Lemma 4.1.** In a steady state \( c, i, s, y, \lambda, \gamma \) and \( \omega \) are constant.

**Proof.** Recall that \( A'(0) = A(0) = 0 \). Then \( \ddot{c} = \ddot{i} = \ddot{s} = \ddot{y} = 0 \) follows directly from (2.8) through (2.10) and Definition 4.1. Differentiating (3.1) through (3.3) with respect to time establishes \( \ddot{\omega} = \ddot{\lambda} = \ddot{\gamma} = 0 \).

In the sequel \( \bar{x} \) denotes the steady state value of the variable \( x \).

**Theorem 4.1.** The steady state is unique on \( \Gamma \).
PROOF. Using Lemma 4.1 and equations (3.1) through (3.3) it follows that (3.4) and (3.5) can be written as

\begin{align}
(4.1) & \quad (\rho + n)b - (a + b) y_k(\bar{p}, \bar{k}) = 0, \\
(4.2) & \quad \rho + m - (a + b) y_p(\bar{p}, \bar{k}) = 0.
\end{align}

Given A.1 it follows that the Jacobian of this system is positive definite and thus it is a $P$-matrix. Hence, by Theorem 4 of Gale and Nikaido (1965) the map defined by the left-hand sides of equations (4.1) and (4.2) for $(\rho, k) \in \Gamma$ is globally univalent. So $\bar{p}$ and $\bar{k}$ are unique. Uniqueness of $\bar{c}$, $\bar{t}$, $\bar{s}$ is then straightforward. \qed

Equations (4.1) and (4.2) are extensions of the neoclassical "golden rule." The marginal product of physical capital should, after correction for the learning by doing and schooling coefficients, equal the rate of time preference and depreciation (recall $\rho + n = \rho' + \mu$). The same holds, necessary changes being made, for the marginal product of production experience.

**Theorem 4.2.** The per capita steady state levels of consumption, capital and production experience are declining functions of the rates of depreciation, net forgetfulness and time preference; increasing functions of the learning by doing coefficient; the population growth rate affects capital and production experience positively, but the effect on consumption is indeterminate. The effect of an increase in the schooling coefficient is ambiguous.

**Proof.** Differentiation of equations (4.1) and (4.2) gives the desired results for $\bar{p}$ and $\bar{k}$. Differentiating equations (2.8) through (2.10) and using the above, yields

\begin{equation}
(4.3) \quad d\bar{c} = \rho d\bar{k} + \frac{1}{b} \rho d\bar{p} - \bar{k} d\rho - \frac{1}{b} \bar{p} d\rho + \frac{1}{b} \bar{y} d\alpha + \frac{1}{b} \bar{s} db .
\end{equation}

The consumption effects follow directly from (4.3). \qed

These results are intuitively appealing. If the costs of capital or production experience go up through an increase in depreciation, or net forgetfulness, or if consumers become less patient, the capital and production knowledge levels decline. An increase in productivity through learning by doing, increases $\bar{p}$ and $\bar{k}$. The ambiguity of schooling expenditures is readily explained by the fact that schooling is costly; thus the marginal benefits of extra schooling have to be weighed against the extra costs, while learning by doing is a positive externality. An example may clarify the situation. Suppose that $y$ is of the Cobb Douglas type

\begin{equation}
(4.4) \quad y = p^{\phi} k^\beta, \quad \text{where} \quad \phi + \beta < 1 .
\end{equation}

Using the above framework the following inequalities can be derived in the steady state

\begin{equation}
(4.5) \quad \partial k / \partial b \leq 0 \quad \text{as} \quad \phi / (1 - \phi) \leq a / b .
\end{equation}
and

$$\frac{\partial p}{\partial b} \leq 0 \quad \text{as} \quad (1 - \beta)/\beta \leq a/b.$$ (4.6)

Three types of regimes can be discerned. For low returns to learning by doing, i.e. "a" low, an increase in the marginal efficiency of schooling expenditures $b$ raises both capital $k$ and production knowledge $p$. The converse is true for high levels of $a$, when schooling is a relatively inefficient means of learning. There is an intermediate range where $p$ (or $k$) rises and $k$ (or $p$) falls as $(1 - \beta)/\beta > \phi/(1 - \phi)$(or $<$. Thus a rise in schooling efficiency decreases or increases the steady state capital stock and production knowledge depending on whether learning by doing is a relatively effective or an ineffective alternative means for production knowledge accumulation.

It is of interest to investigate the endogeneity of schooling and learning by doing for the steady state welfare level. Consider two economies that are identical except for the aspect of human capital accumulation. The first economy possesses the schooling and learning by doing possibilities for production experience accumulation which were described above, whereas the production experience level is exogenously given to the second economy. How do the two economies compare with respect to their steady state values? To make a fair comparison the second economy is given the first economy’s steady state production experience $\bar{p}$. Its steady state capital stock $k^*$ can be derived from the well known rule

$$y_k(\bar{p}, k^*) = \rho + n.$$ Since by (4.1)

$$y_k(\bar{p}, k^*) = \rho + n > (\rho + n) \frac{b}{a + b} = y_k(\bar{p}, \bar{k}),$$

we have $k^* < \bar{k}$. So the second economy has less capital in the steady state and therefore its aggregate output is less. This does not imply, however, that aggregate consumption is also lower. Consider e.g. the case with almost no learning by doing, i.e. as $a \to 0$ then $k^* \to \bar{k}$, but because schooling is costly, it follows that $\bar{c} < c^*$. The economy possessing the schooling technology will have to invest in production knowledge through schooling whereas the other economy does not. If the outlays on schooling are low or if the rate of net forgetfulness is small, one would expect the economy possessing the schooling technology to be better off. To illustrate this consider again the Cobb Douglas case, i.e. equation (4.4).

Use upper bars and stars to denote the steady state values of the variables from the first and the second economy respectively. Some straightforward but tedious calculations give

$$\bar{c} = \frac{a + b}{b} \left[ 1 - \frac{n\beta}{\rho + n} - \frac{m\phi}{\rho + m} \right] \bar{y},$$

$$p^* = \bar{p}, \quad k^* = \left(\frac{a + b}{b}\right)^{1/(\beta - 1)} \bar{k},$$
\[ c^* = \left( \frac{a + b}{b} \right)^{\beta(\beta - 1)} \left[ 1 - \frac{n\beta}{\rho + n} \right] \tilde{y}. \]

Thus the economy with endogenous human capital accumulation is better off if and only if

\[ 1 - \frac{n\beta}{\rho + n} - \frac{m\phi}{\rho + m} > \left( \frac{b}{a + b} \right)^{1/(1 - \beta)} \left[ 1 - \frac{n\beta}{\rho + n} \right]. \]

So, if the rate of net forgetfulness is small, the economy with endogenous human capital accumulation is better off. Because learning by doing is a positive externality to production, a reduction in the coefficient \( a \), holding \( b \) fixed, may tip the balance to the other side. Similarly, if schooling is a relatively efficient means of experience accumulation, i.e. \( b/a \) is high, and because schooling is costly as it reduces the feasible consumption set for the economy with endogenous learning, the economy with the fixed experience level may be better off.

5. STABILITY AND EXISTENCE

The aim of this section is to investigate the local and global stability properties of the steady state of our system. Control problems with more than one state variable are generally hard to handle analytically, see Pitchford (1977). In our case the problems are surmountable.

Define \( x = (x_1, x_2, x_3, x_4) \) as follows: \( x_1 = k - \bar{k}, x_2 = \rho - \bar{\rho}, x_3 = \lambda - \bar{\lambda}, x_4 = \gamma - \bar{\gamma} \). Since \( i = i(x) \) and \( s = s(x) \) by Lemma 3.1, the system (2.8), (2.9), (3.4), (3.5) can be rewritten as (using (3.3) as well):

\[ \dot{x}_1 = i(x) - n[x_1 + \bar{k}]. \]
\[ \dot{x}_2 = ay(\bar{\rho} + x_2, \bar{k} + x_1) - A(i(x) - n[x_1 + \bar{k}]) + bs(x) - m[x_2 + \bar{\rho}], \]
\[ \dot{x}_3 = [\rho + n][x_3 + \bar{\lambda}] - [x_4 + \bar{\gamma}][a + b] y_k(x_2 + \bar{\rho}, x_1 + \bar{k}) \]
\[ \quad - n[x_4 + \bar{\gamma}] A'(i(x) - n[x_1 + \bar{k}]), \]
\[ \dot{x}_4 = [\rho + m][x_4 + \bar{\gamma}] - [x_4 + \bar{\gamma}][a + b] y_p(\bar{\rho} + x_2, \bar{k} + x_1), \]

or, in shorthand,

\[ \dot{x} = g(x). \]

Let \( x(t; x(0)) \) be defined as the solution of (5.5) when the initial state is \( x(0) \).

Stability is investigated according to the following two definitions.

**Definition 5.1.** The steady state \( x = 0 \) is locally asymptotically stable (LAS) if there exists some open neighborhood of \((0, 0)\) such that for all \((x_1(0), x_2(0))\) in this neighborhood there exists \((x_3(0), x_4(0))\) such that \( x(t; x(0)) \rightarrow 0 \) as \( t \rightarrow \infty \).
DEFINITION 5.2. The steady state \( x = 0 \) is globally asymptotically stable (GAS) if for all \((k(0), p(0)) \in \text{int}(\Gamma)\) there exists \((x_3(0), x_4(0))\) such that \(x(t, x(0)) \rightarrow 0\) as \(t \rightarrow \infty\).

Local asymptotic stability is dealt with first.

THEOREM 5.1. The steady state is LAS.

PROOF. What we wish to show is that \( x = 0 \) is a regular saddlepoint. To this end we first show that there exists a two dimensional manifold \( M \), containing the origin, such that for any \( x(0) \) on \( M \) \( x(t, x(0)) \rightarrow 0 \). Coddington and Levinson (1955, Ch. 13, Th. 4.1) prove that such a \( M \) exists if \( g(x) \) can be written as \( Ax + f(x) \) where \( A \) is a constant matrix with two eigenvalues that have negative real parts and two that have positive real parts, \( f(x) \) is continuous for small \( x \), and \( f'(x) = o(1) \) as \( \|x\| \rightarrow 0 \). Define \( f(x) = g(x) - g'(0)x \). Then, \( g(x) = g'(0)x + f(x) \). Evidently

\[
f'(x) = g'(x) - g'(0) \rightarrow 0 \quad \text{as} \quad \|x\| \rightarrow 0.
\]

In the Appendix, part 2, it is shown that \( g'(0) \) has two positive and two negative real eigenvalues. Hence, there exists a stable two dimensional manifold \( M \). Even though \( M \) exists, it may not be possible to find an open neighborhood of \((0, 0)\) such that for all \((x_1(0), x_2(0))\) in this neighborhood there is an \((x_3(0), x_4(0))\) on \( M \). To conclude the proof, we therefore show in the Appendix, part 3, that the desired open neighborhood does exist. \(\square\)

The main implication of Theorem 5.1 is that it shows that the problem (2.7) through (2.10) has a solution in some small neighborhood of \( x = 0 \), compare Theorem 3.1. To appreciate the last part of the proof, note that in control problems with one state variable it is usually straightforward to verify that the stable branch is not orthogonal to the state space.

Global stability of a solution holds as well if the discount factor is sufficiently small.

THEOREM 5.2. The steady state is GAS provided \( \rho \) is "small."

PROOF. By Theorem 3.2 of Brock and Scheinkman (1976) \( x = 0 \) is globally asymptotically stable if the "curvature matrix" \( Q \) is positive definite and two additional conditions concerning the trajectories are satisfied as well. The matrix \( Q \) is defined as

\[
Q = \begin{bmatrix}
H_{11} & \frac{1}{2} \rho I \\
\frac{1}{2} \rho I & -H_{22}
\end{bmatrix}
\]

where \( H_{11} \) and \( H_{22} \) are given by

\[
Q = \begin{bmatrix}
H_{11} & \frac{1}{2} \rho I \\
\frac{1}{2} \rho I & -H_{22}
\end{bmatrix}
\]
\[ H_{11} = \begin{bmatrix} \frac{1}{\gamma A''} & -\frac{A' + b}{\gamma A''} \\ -\frac{A' + b}{\gamma A''} & \frac{(A' + b)^2}{\gamma A''} - \frac{b^2}{U''} \end{bmatrix}, \quad H_{22} = \gamma(a + b) \begin{bmatrix} y_{kk} & y_{kp} \\ y_{kp} & y_{pp} \end{bmatrix}, \]

and \( I \) is the \( 2 \times 2 \) identity matrix. As the diagonal elements and determinants of \( H_{11} \) and \(-H_{22}\) are positive (by A.1 through A.3), these submatrices are positive definite. Therefore, for any \((k, p) \in \text{int}(\Gamma)\), there exists a \( \rho \) sufficiently small such that \( Q \) is positive definite. In addition the following two conditions are needed. First, we want the value function to be differentiable. It is easily verified that the conditions for this to hold as stated in Benveniste and Scheinkman (1979) are satisfied. Second, the transversality conditions (3.6) and (3.7) do hold. It follows from Brock (1977, fn. 12) that \( x = 0 \) is globally asymptotically stable.

6. LOCALIZED PRODUCTION EXPERIENCE AND SCHOOLING

In the steady state the capital-labor ratio is constant, hence there can be no loss in production experience due to learning by doing cum adjustment costs. The effect of localized production experience manifests itself only on the path toward the steady state. Therefore the interesting question is: what is the effect of an increase in the adjustment costs on the speed of convergence? In Section 5 it was shown that the matrix \( g'(0) \) has four distinct and real eigenvalues, say \( \phi_1, \ldots, \phi_4 \), such that \( \phi_1 < \phi_2 < 0 < \phi_3 < \phi_4 \). The time to stationarity along the stable manifold in a neighborhood of the steady state is dominated by \( \phi_2 \). Suppose we parameterize the adjustment cost function \( A \) by introducing the parameter \( \xi \):

\[ \xi A(i - nk). \]

An increase in \( \xi \) then signifies an increase in the adjustment costs. The effect this has on the characteristic polynomial is simply to replace \( A'' \) by \( \xi A'' \). Hence one might, instead of this parametric approach, just as well look at the effect a change in \( A'' \) has on \( \phi_2 \). Thus identify an increase in \( A'' \) with a parametric increase in the adjustment costs. Similarly, identify an increase in \( U'' \) with a parametric increase in felicity.

**Theorem 6.1.** A parametric increase in adjustment costs decreases the local speed of convergence in a neighborhood of the steady state.

**Proof.** From the Appendix, part 2, we know the characteristic polynomial reads

\[ \psi(\phi; A'') = \phi^4 - 2 \rho \phi^3 + [\rho^2 - \tau(A'') - \beta] \phi^2 + \]
\[ + \rho [\tau(A'') + \beta] \phi + \tau(A''), \]

where \( \beta, \tau \) and \( \pi \) are positive, and \( \psi \) has real roots \( \phi_1 < \phi_2 < 0 < \phi_3 < \phi_4 \). By the implicit function theorem
\[ \partial \phi_2 / \partial A'' = -[\partial \psi(\phi_2; A'') / \partial A''] / [\partial \psi(\phi_2; A'') / \partial \phi_2]. \]

The denominator is positive as may be seen from the slope of the graph of \( \psi \) at \( \phi_2 \) in Figure 2, see the Appendix, part 2. Hence \( \partial \phi_2 / \partial A'' \) is positive if and only if, the numerator is negative. The numerator can be written as:

\[ \partial \psi(\phi_2; A'') / \partial A'' = \frac{1}{A''} \left[ \tau \phi_2^2 - \rho \phi_2 - \pi \right]. \]

From the function \( f(\phi) \) introduced in the Appendix, part 2, it is immediate that the term in between the square brackets is negative.

Two opposing forces are at work in the two state variable model. To illustrate this, consider first the one sector neoclassical growth model with adjustment costs:

\[ k = i - nk - A(i - nk), \]
\[ y(k) = c + i. \]

The characteristic polynomial of the linearized system reads

\[ \psi(\phi; A'') = \phi^2 - \rho \phi - \sigma(A''), \]

where

\[ \sigma(A'') = U' y_{kk} / (U'' - U'A'') > 0. \]

Clearly \( \psi(0; A'') = -\sigma(A'') < 0 \) and \( \partial \psi(0; A'') / \partial \phi = -\rho \). The situation is depicted in Figure 1. Let \( \phi_1 < 0 < \phi_2 \), and \( \psi(\phi_1) = \psi(\phi_2) = 0 \). It follows that \( \partial \psi(\phi_1; A'') / \partial \phi_1 < 0 \), and \( \partial \psi(\phi_1; A'') / \partial A'' = -\partial \sigma(A'') / \partial A'' > 0 \). Hence, \( \partial \phi_1 / \partial A'' > 0 \). This is caused by an increase in \(-\sigma(A'')\), while the slope remains constant at \(0, -\sigma\).

In the two state variable model \( \partial \psi(\phi_2; A'') / \partial \phi_2 > 0 \), but this time there is not only a reduction in \( \psi(0; A'') \), but also a spillover effect from one "sector" to the other, i.e., a change in the slope of \( \partial \psi(0, A'') / \partial \phi \). This spillover effect tends to increase the convergence speed, but, as Theorem 6.1 shows, is dominated by the "one sector" effect.

The local speed of convergence depends on other factors as well; the results are given in Theorem 6.2.

**Theorem 6.2.** The following will reduce the local speed of convergence:

(i) an increase in felicity,
(ii) an increase in the rate of time preference,
(iii) a decrease in the learning by doing coefficient.

The effect of a change in the schooling coefficient is, however, ambiguous.

---

6 For a proof of Theorem 6.2, which is similar to the proof of Theorem 6.1, the interested reader is referred to van Marrewijk (1988).
The results are intuitively appealing. If agents are more patient, for instance, one expects the rate of convergence to increase because the agents are willing to defer some consumption and increase investment or schooling. The ambiguity of the schooling effect $b$ derives from the presence of the factors $2b^2 y_{kp}$ and $-b^2 y_{pp}$ in the coefficient $\tau$ of the characteristic polynomial (6.2). Apart from these two factors
all effects go into the same direction as changes in the learning by doing coefficient $a$, as expected on the basis of equation (2.4). But because schooling lays a claim on the budget, while learning by doing is a positive externality to production, it may not pay to increase the accumulation speed as it retards consumption during the transition phase.

7. CONCLUSION

The paper unifies three factors of economic growth, i.e. physical capital accumulation, schooling and learning by doing, into one growth model. The latter two factors contribute to production experience accumulation. This production experience is specific to the techniques which are in operation such that technique changes induce adjustment costs on part of the human capital accumulation. This localization effect links the three growth factors. The formalization gives rise to a two state variable optimal control problem. The steady state and the path to it are unique and, provided the discount factor is sufficiently low, the steady state is globally asymptotically stable. The results of economic interest concern the steady state welfare effects of human capital accumulation and the speed of convergence to the steady state. Among the first set of results are the decrease in the per capita steady state levels of consumption, production experience and capital following an increase in either the rates of depreciation, forgetfulness, time preference or a decrease in the learning by doing effect. The second set of results shows that the speed of adjustment close to the steady state reduces through increases in either the adjustment costs, the rate of time preference or felicity or a decrease in the learning by doing effect. In both cases the effects of changes in the marginal efficiency of schooling expenditures are ambiguous because schooling is costly, while the learning by doing effect is determinate because it arises as a positive externality from production.

We investigated a closed plan economy and more research into an open economy with private decision makers is needed. The model developed here, however, is the first to incorporate several different aspects of economic growth due to human capital and physical capital accumulation into one model. It can serve as a benchmark for future research into the interaction of schooling, learning by doing, optimal investment and localized growth. In particular, we have generated a set of predictions on learning by doing and schooling expenditures that warrant empirical evaluation.

Erasmus Universiteit Rotterdam, The Netherlands
Katholieke Universiteit Leuven, Belgium
Technische Universiteit Eindhoven, The Netherlands

APPENDIX

Part 1. Comparative Statics. In this part of the appendix we perform a comparative statics analysis for $c$, $i$ and $s$ around the steady state. The set of equations (2.10), (3.1) through (3.3) can be written as
\[-U' + b\gamma = 0,\]
\[\lambda - \gamma A' - U' = 0,\]
\[c + i + s - y(p, k) = 0.\]

Due to Lemma 3.1 we may write \(c = c(k, p, \lambda, \gamma), s = s(k, p, \lambda, \gamma),\) and \(i = i(k, p, \lambda, \gamma).\) Comparative statics yield
\[c_k = 0 \quad s_k = y_k - n \quad i_k = n > 0,\]
\[c_p = 0 \quad s_p = y_p > 0 \quad i_p = 0,\]
\[c_\lambda = 0 \quad s_\lambda = -1/\gamma A'' < 0 \quad i_\lambda = 1/\gamma A'' > 0,\]
\[c_\gamma = b/U'' < 0 \quad s_\gamma = (b + A')/\gamma A'' - b/U'' \quad i_\gamma = -(b + A')/\gamma A''.\]

**Part 2. Eigenvalues.** In Section 5 of the paper we analyze the system of differential equations, \(\dot{x} = g(x).\) Here it is shown that the matrix \(g'(0)\) has two positive and two negative real eigenvalues. Elementary calculations lead to the following \(g'(0)\) matrix, where \(h = a + b,\) and all functions are evaluated at \(x = 0:\)
\[
g'(0) = \begin{bmatrix}
0 & 0 & 1/\gamma A'' & -b/\gamma A'' \\
\rho b & \rho & -b/\gamma A'' & b^2/\gamma A'' - b^2/U'' \\
-\gamma h y_{kk} & -\gamma h y_{pk} & \rho & -\rho b \\
-\gamma h y_{pk} & -\gamma h y_{pp} & 0 & 0
\end{bmatrix}.
\]

Some tedious algebra then gives the characteristic polynomial:
\[
\psi(\phi) = \phi^4 - 2\rho \phi^3 + (\rho^2 - \tau - \beta)\phi^2 + \rho(\tau + \beta)\phi + \pi,
\]

where
\[
\tau = (a + b) \left[2by_{pk} - y_{kk} - b^2 y_{pp}\right]/A'' > 0,
\]
\[
\beta = \gamma b^2 (a + b)y_{pp}/U'' > 0,
\]
\[
\pi = -\gamma b^2 (a + b)^2 \left[y_{kk} y_{pp} - y_{pk}^2\right]/A'' U'' > 0.
\]

The inequalities directly follow from A.1, A.2 and A.3. We now show that \(\psi(\phi) = 0\) has four real roots, two of them being positive and two being negative. First note that \(\psi(0) = \pi > 0.\) Consider \(f : \mathbb{R} \to \mathbb{R}\) defined by
\[
f(\phi) = \tau \phi^2 - \rho \tau \phi - \pi.
\]

Clearly there exist real \(\phi_5\) and \(\phi_6\) such that \(\phi_5 < 0 < \phi_6\) and \(f(\phi_5) = f(\phi_6) = 0.\) Note that for \(i = 5,6: \phi_i^2 = \rho \phi_i + \pi/\tau, \phi_i^3 = \rho \phi_i^2 + (\pi/\tau)\phi_i, \phi_i^4 = \rho \phi_i^3 + (\pi/\tau)\phi_i^2.\) Using this, it follows that
psi(\phi_5) = psi(\phi_6) = \frac{\pi}{\tau^2} (\pi - \beta \tau) < 0 ,

because (\pi - \beta \tau) = -b^2 (a + b)^2 \gamma [2b \gamma_{pk} \gamma_{pp} - b^2 \gamma_{pp}] / A'' U'' < 0.

Clearly, see Figure 2, there exist real \phi_1, \phi_2, \phi_3 and \phi_4 such that \phi_1 < \phi_5 < \phi_2 < 0 < \phi_3 < \phi_6 < \phi_4 and psi(\phi_i) = 0 for i = 1, ..., 4.

**Part 3. A Surjective Map.** We show that the two dimensional manifold M of Theorem 5.1 is not orthogonal to the (x_1, x_2) plane. Loosely speaking this ensures that (x_1, x_2) will change as (x_3, x_4) change, because otherwise for some (x_1(0), x_2(0)) in this neighborhood there may be no pair (x_3(0), x_4(0)) on M. Write g'(0), which is given in part 2, by

\[ G = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix} , \]

where G_i is a 2 x 2 matrix, i = 1, 2, 3, 4. Note that G is square but not symmetric. Given that all four eigenvalues are distinct, it follows by result vii a in Rao (1973, p. 43) that there exists a nonsingular matrix P such that \[ P^{-1} G P = D , \]

where D is the diagonal matrix with the eigenvalues of G on the diagonal and the negative
eigenvalues are in the upper left-hand corner while the positive eigenvalues are in the lower right-hand corner, and $P$ is the matrix with the corresponding eigenvectors of $G$. Partition $P$ as follows

$$
P = \begin{bmatrix}
P_1 & P_2 \\
P_3 & P_4
\end{bmatrix}.
$$

Thus $[P_2^T P_4^T]$ is the matrix with eigenvectors that are associated with the positive eigenvalues.

Define $y = P^{-1}x$. According to Theorem 4.1 in Coddington and Levinson (1955, ch. 13) there are continuous functions $h_3(y_1, y_2)$ and $h_4(y_1, y_2)$ defined for small $|y_1|$ and $|y_2|$ such that $y_3 = h_3(y_1, y_2)$ and $y_4 = h_4(y_1, y_2)$ constitute a two dimensional manifold in $y$ space. Our two dimensional manifold $M$ in $x$ space is now obtained as

$$
x = P(y_1, y_2, h_3(y_1, y_2), h_4(y_1, y_2))^T.
$$

In the proof of Theorem 5.1 it was shown that $f(x) = g(x) - g'(0)x$ is differentiable. Therefore, see Coddington and Levinson (1955, Th. 4.2, ch. 13), $\partial h_j/\partial y_i = 0$ at $y_j = 0$ for $j = 3, 4$ and $i = 1, 2$. Thus to ensure that the two dimensional manifold $M$ is well defined for all state variables $(x_1, x_2)$ in some neighborhood of $(0, 0)$ it is sufficient to show that the upper left-hand $2 \times 2$ submatrix $P_1$ of $P$ is nonsingular. In that case we are ensured that variations in $(y_1, y_2)$ imply two dimensional variations in $(x_1, x_2)$.

Straightforward but tedious calculations show that any eigenvector $e = (w, x, y, z)$ satisfies

$$(w, x, y, z) = (e_1(\phi), x, y, z),$$

where $e_1$ is a well defined function of the eigenvalues $\phi$. (To achieve this manipulate the first, third and fourth equations of the system $Ge = \phi e$, and show that the denominator of $e_1$ is nonzero for any negative eigenvalue $\phi$.) Recall $\phi_1, \phi_2$ are the two negative eigenvalues. Hence $P_1$ is nonsingular if

$$
\det P_1 = x^2[e_1(\phi_1) - e_1(\phi_2)]
$$

is nonzero. Evidently, by the above expression for $e$, $x$ cannot be zero as $P$ is nonsingular. Moreover, after some tedious but simple calculations it can be shown that $e_1(\phi_1) = e_1(\phi_2)$ cannot hold when $\phi < 0$.

REFERENCES


