Modeling Reasons for Firms' Growth

Matti Estola*
University of Joensuu, Dept of Business and Economics,
P.O. Box 111, FIN-80101 Joensuu, Finland.
Tel: +358-13-2514587, Fax: +358-13-2513290,
E-mail: matti.estola@joensuu.fi

21st October 2003

Abstract

Static neoclassical framework cannot be applied to model time dependent processes or increasing returns to scale in production. The dynamization of the neoclassical theory of a firm by dynamic optimization, on the other hand, assumes inconsistent profit functions with the former. As a solution to these problems, we present a dynamic theory of a firm consistent with the static neoclassical theory. Possible reasons for a firm's growth are an increase in demand of the firm's products, a decrease in costs or increasing returns to scale in production. We model the reasons causing these elements in firms' production processes and present block diagrams of the models to reveal their control theoretic nature. (JEL B41, D21, D24)

Keywords: Dynamics, firms' growth, control theory.

1 Introduction

According to Mirowski (1989a), neoclassical economic thinking is based on two distinct elements: egoistic economic agents by Smith (utility maximizing consumers by Jevons, Menger and Walras) and the mathematical metaphor of classical mechanics. The latter can be understood by the progenitors of neoclassical economics who were engineer level physicists. Concept equilibrium was borrowed from physics and introduced in economics by Canard at

*Support from the Yrjö Jahnsson Foundation (grant # 4630) is gratefully acknowledged.

1801 (Mirowski 1989b). Although equilibrium is ‘a balance of forces’ situation, in economics the balancing ‘forces’ have not been defined. In order to understand the adjustment process, however, we should define the forces which ‘push’ economic quantities toward their equilibrium values in a stable case, or cause their evolution with time in a non-stable case.

The existence of forces acting upon economic quantities can be argued indirectly; every changing quantity (price, wage, exchange rate etc.) tells the existence of reasons (forces) causing these changes. This is analogous with arguing the existence of the gravitational force field by dropping a pen; without the force field the pen would not move. Fisher (1983 pp. 9-12) writes: “... I now briefly consider the features that a proper theory of disequilibrium adjustment should have ... if we are to show under what conditions the rational behavior of individual agents drives an economy to equilibrium. ... Such a theory must involve dynamics with adjustment to disequilibrium over time modeled. ...the most satisfactory situation would be one in which the equations of motion of the system permitted an explicit solution with the values of all the variables given as specific, known functions of time. ... Unfortunately, such a closed-form solution is far too much to hope for. ...the theory of the household and the firm must be reformulated and extended where necessary to allow agents to perceive that the economy is not in equilibrium and to act on that perception. ... A convergence theory that is to provide a satisfactory underpinning for equilibrium analysis must be a theory in which the adjustments to disequilibrium made by agents are made optimally.”

According to Arrow & Hahn (1971, p. 12), the accepted definition for market stability is that of Samuelson (1942) who writes: “In the history of mechanics, the theory of statics was developed before the dynamical problem was even formulated. But the problem of stability of equilibrium cannot be discussed except with reference to dynamical considerations ... we must first develop a theory of dynamics.” Mas-Colell et al. (1995, p. 620) write: ”A characteristic feature that distinguishes economics from other scientific fields is that, for us, the equations of equilibrium constitute the center of our discipline. Other sciences, such as physics or even ecology, put comparatively more emphasis on the determination of dynamic laws of change. The reason, informally speaking, is that economists are good (or so we hope) at recognizing a state of equilibrium but are poor at predicting precisely how an economy in disequilibrium will evolve. Certainly there are intuitive dynamic principles: if demand is larger than supply then price will increase, if price is larger than marginal costs then production will expand...”

We base our modeling on the principles stated by these authors. We define the forces which act upon the production of a firm, and use these forces to model the dynamics of production of the firm in real time. The possible asymptotic steady-state of the firm is the neoclassical one: marginal revenues equals marginal costs for every input, and the equilibrium use of every input maximizes the firm's profit. To define the 'economic forces acting upon the use of inputs of a firm', we assume that the firm's managers like to better the firm's situation when possible, as Fisher stated above. We believe that the *'economic agents' desire to better their situation'* is the cause for the observed dynamics in economies.

Dynamic control theoretic models concerning firms' behavior completely base on dynamic optimization techniques; see for example Bensoussan et al. (1974) or Chiang (1992). However, it has been shown in physics that Hamilton's and Newton's modeling principles produce the same equations of motion, and Newton's method assumes only differential equations while Hamilton's principle requires dynamic optimization. Hamilton's principle can be stated as: *"The actual path a dynamic system follows is that which minimizes the time integral of the difference between kinetic and potential energies of the moving body"* (Marion & Thornton 1988 p. 192). Newton's principle, on the other hand, relies on the forces acting upon the moving object, and can be stated as: *"A body acted upon by a force moves in such a manner that the (time) rate of change of momentum equals the force"* (ibid. p. 44). Due to this relative simplicity, all practical engineering control applications rely on Newton's principle. In engineering control models, it is also common to present a block diagram of the system before its behavior is analyzed mathematically, see Marshall (1978) or Ogata (1997).

We propose a similar simplification for control theoretic modeling in economics as Newton's principle offers in physics. The 'force vector', which causes the dynamics of the use of every input of a firm, guides the firm with time toward its profit maximizing state, and so dynamic optimization techniques are not needed. We also present various block diagrams of firms' production processes to reveal their control theoretic nature. The Newtonian type of 'dynamic law for the firm's adjustment of its inputs' is used as a building block in our control systems, as is common in engineering. By the models we study reasons for production dynamics at firm level.

Modern theory of endogenous growth by Romer, Lucas and others arose as a critique to Solow's (1956) explanation of economic growth by exogenous technological development. Romerian (1990) theory models the macro-level growth of an economy by endogenous development in technology due to firms' investments in human capital. Romer classifies human capital as partly rival and partly non-rival, i.e., of public good in nature. For example, any new technology contained in a durable good becomes public when these goods are sold. This way technology diffuses throughout the firms in an economy and around the world. In Romer's theory, the increase in the number of existing producer durables increases the efficiency of production processes and this way promotes growth.

Lucas (1988) measures an individual's 'human capital' by his level of skills. The economy-level human capital then has both rival and non-rival components as in Romer's model; an improvement in skills of every individual increases the average level of skills in the economy. Lucas stresses the role of people's decisions how to allocate their time between working and developing their skills in the accumulation of human capital. Although Romer and Lucas both claim that the development of human capital is, as Lucas it states, the *'unobservable magnitude or force'* behind economic growth, by human capital they obviously mean different things.

Common to Romer's and Lucas' theories is that they model the balanced growth of an economy. When we measure the growth of an economy, however, we always observe varying growth rates at industries. Every growing economy contains industries which create the growth and those benefiting from the growing ones in the form of producing input goods or services to them; industries with negative growth rates may also exist. In a growing economy, the structure of consumption changes with consumers' wealth according to their preferences. These remarks question the tradition to model growth in an aggregated framework and the concentration on balanced growth.

If we accept that the economy-level growth has a micro foundation at the firm level, and we still accept that depending on the level of wealth of an economy its consumers allocate their funds to different kind of goods, we see that at various states of the development of an economy different industries have a potential to grow fast (we have no evidence that the needs of human beings were limited). In an agricultural economy, technological development in fielding and farming may be the most important source of economic growth, and in a 'post-manufactured' economy, the development of information technology may be most important. Relatively poor economies producing export goods to wealthier economies may also grow fast. Those industries, which may create export led growth, vary with the relative wealth and technology levels of economies. The process of economic development is the same in all these cases, although those industries and firms which create the growth vary. The theories of economic development and growth thus model different stages of the same process.

The reasons behind technological development in an economy vary between researches. The origin of Solow's type of 'exogenous technological development' may be the 'learning by doing' of employees at work or normal technological progress in producer durables with time. In Romer's model, a research sector creates new designs of producer durables for firms which produce consumption goods, and these firms use the producer durables in an optimal way. New desings and durable goods based on them are produced due to the profit-seeking interest of their producers. Lucas, on the other hand, assumes that an individual worker is paid according to his marginal productivity, which gives him a rationale to develop his skills.

The reasons which increase the efficiency of labor may be exogenous or endogenous for a firm. Exogenous reasons are learning by doing at work — which though takes place in a firm but does not require an investment in it — and better education or technological level in the society. Endogenous reasons are an increase in the number of capital goods or more effective ones available for workers, or firms' investments in production methods and employees' skills; the last two can be included in the category of 'human capital'. We model the role of these elements in the growth of firms' productions and show that the 'unobservable human capital' is not a necessity for economic growth. Every factor, which increases the productivity of inputs or reduces the costs of a firm, creates the firm a potential for growth.

2 Kinematics of Production

In the following we study the behaviour of a single firm and denote by $q(t)$ $(unit/y)$ its velocity of production at time moment t[1]. Unit y may be one day, week, month or other suitable time unit depending on the firm. The production process is assumed continuous, although it can be seized at any time. If a firm operates 8 hours per day with one lunch and two coffee breaks, we can still consider the production process continuous by assuming that resting at night and eating during work-days are essential elements of the process. Finished products, like radio sets, bits of sugar etc., are measured in discrete units, for instance, by $numbers$, $kg's$, $liters$ or any other quantity units; $q(t)$ may thus take values 3/8 $(kg/hour)$ or 3.3 $(unit/day)$ etc.

The time path of accumulated production of a firm is analogous to that of the mileomiter of a car registering full kilometers in discrete units, and parts of a kilometer every second when running. The speed of a car varies when it moves and is zero when staying at traffic lights, gasoline station, in garage etc. If 4 radio sets have been completed in 7.15 hours in a factory, the average velocity of production during the period is $4/7.15$ $(unit/hour)$. We can express this as $(4 \times 24)/7.15$ $(unit/day)$ if one work-day consists of 24 hours. If one work-day consists of 8 hours, the velocity is $(4 \times 8)/7.15$ $(unit/day)$. The measured velocity of production thus depends on the length of one work-day. This is analogous to a car having driven 200 kilometers in

[1]Measurement units are denoted in braces after the quantities. A system of measurement units for economics is given in de Jong (1967).

2.5 hours with the driver having had one lunch break of 30 minutes during the trip. The average speed during the trip was then 200/2.5 (km/h), but on the road the average speed was 100 (km/h).

The accumulated production of the firm till time moment t (the accumulated kilometers a car has driven) denoted by $Q(t)$ and measured in units $unit$ (a marginal change in time ds is measured in units y) is

$$Q(t) = Q(t_1) + \int_{t_1}^{t} q(s)ds, \quad Q'(t) = q(t), \quad Q''(t) = q'(t),$$

where $Q(t_1)$ is the accumulated production of the firm from its foundation till moment t_1, $Q'(t) = q(t)$ ($unit/y$) the momentous velocity of production and $Q''(t) = q'(t)$ ($unit/y^2$) the momentous acceleration of production at moment t. This kinematics of production is a necessary prelude for production dynamics analogous to Newtonian mechanics.

3 Dynamisation of Neoclassical Theory

A common way to transform the neoclassical theory of a firm into dynamic form is to assume that firms maximize their expected profits from a certain time period. We simplify the situation by omitting uncertainty and assume the profit function for a firm for time unit y as $\Pi(q(t), t) = P(t)q(t) - C(q(t))$, where $P(t)$ ($eur/unit$) is the unit price of the product of the firm, $q(t)$ ($unit/y$) the velocity of production and $C(q(t))$ (eur/y) the cost function. The dynamic optimization problem concerning the current value of the firm's profit from time period (t_0, t_1) is

$$\max_{q(t)} \int_{t_0}^{t_1} F\big(q(t), t\big)dt = \max_{q(t)} \int_{t_0}^{t_1} e^{-rt}\Pi\big(q(t), t\big)dt, \tag{1}$$

where e^{-rt} the discount factor with constant interest rate r ($1/y$); see the Appendix Part A. The necessary condition for (1) — together with possible boundary conditions — is the Euler -equation

$$\frac{\partial F}{\partial q(t)} - \frac{d}{dt}\left(\frac{\partial F}{\partial q'(t)}\right) = 0 \;\Rightarrow\; e^{-rt}\frac{\partial \Pi}{\partial q} = 0 \;\Rightarrow\; \frac{\partial \Pi}{\partial q} = 0.$$

The necessary condition for problem (1) thus equals with that of static neoclassical theory. Dynamic optimization was used, however, to get an equation of motion for $q(t)$. The above implies that for this function $\Pi(.)$ should depend on $q'(t)$, but $q'(t)$ does not exist in profit functions in the literature of static neoclassical theory. Static neoclassical theory and its dynamisation by dynamic optimization thus assume inconsistent profit functions, see Chiang (1992 pp. 49, 69, 292). Another minor shortcoming in the existing dynamic models of firm behavior is that they do not actually model the dynamics of production, but the evolution of the unit price or the accumulation of the capital of a firm. The link from price and capital to production can though be found in these models, see also Dixit and Pindyck (1994 pp. 254, 359). Due to these inconsistencies, we introduce a dynamic extension for the static neoclassical theory consistent with it.

4 The Production Process of a Firm

We study the production process of a one-product firm using labor, capital goods and human capital as inputs. The production function is $q = f(L, K, R, t)$, where q $(unit/y)$ is the velocity of production of the firm, L (h/y) the use of labor of the firm in its ordinary production process, R (h/y) the use of labor of the firm in its production of human capital, K (eur) the monetary value of the physical capital of the firm and time t represents possible exogenous technological development. The partial derivatives of the production function with respect to all variables are assumed positive.

If the capital of the firm is financed by loans, the costs of capital are the interest costs rK (eur/y) where the average interest rate r $(1/y)$ of the loans is determined at the financial market. If, on the other hand, capital is financed by shares, the interest costs of the capital rK are paid to the share holders (we assume an equal interest rate in both cases). For simplicity, capital goods are assumed not to deteriorate and we omit the redemption of the loan capital; we can thus think that the capital is financed by shares. Labor is assumed homogeneous so that every employee in the production of final goods and human capital, respectively, is as productive. For simplicity, perfect competition is assumed in the two labor markets so that the firm has no power concerning the wages w_1, w_2 (eur/h) for both types of labor.

The profit of the firm during time unit y is then

$$\Pi(t) = P(t)f(L(t), K(t), R(t), t) - (1 + \tau)[w_1(t)L(t) + w_2(t)R(t)] - r(t)K(t),$$

where τ is fixed assumed wage tax rate. The after tax income of both types of labor are w_1L and w_2R and the firm pays $\tau[w_1L + w_2R]$ (eur/y) taxes to the government. The time derivative of the profit is

$$\begin{aligned}\Pi'(t) &= \frac{\partial\Pi}{\partial P}P'(t) + \frac{\partial\Pi}{\partial L}L'(t) + \frac{\partial\Pi}{\partial K}K'(t) + \frac{\partial\Pi}{\partial R}R'(t) + \frac{\partial\Pi}{\partial w_1}w_1'(t) \\ &+ \frac{\partial\Pi}{\partial w_2}w_2'(t) + \frac{\partial\Pi}{\partial r}r'(t) + \frac{\partial\Pi}{\partial t}. \qquad (2)\end{aligned}$$

Eq. (2) shows that changes in the two wages and interest rate affect the profit, but these variables are beyond control of the managers of the firm. The partial derivative $\partial\Pi/\partial t$ shows that elements like workers' learning, better education in the society etc. — which take place gradually with time and create no costs for the firm — positively affect the firm's profitability. The variables, which the firm's managers can affect, are the unit price, the two labor inputs and physical capital. The assumed adjustment rules are:

$$\begin{aligned} x'(t) &> 0 \quad \text{if} \quad \frac{\partial\Pi}{\partial x} > 0, \\ x'(t) &< 0 \quad \text{if} \quad \frac{\partial\Pi}{\partial x} < 0, \\ x'(t) &= 0 \quad \text{if} \quad \frac{\partial\Pi}{\partial x} = 0, \quad x = P, L, K, R. \end{aligned}$$

These rules make the first four additive terms on the right hand side of Eq. (2) non-negative, i.e. by these adjustments the managers increase the firm's profit with time. The last rule corresponds to a steady state situation where the adjustment of the variable does not contribute to profit. The proposed adjustment rules are in line with neoclassical theory which corresponds to situation $\partial\Pi/\partial x = 0$, $x = P, L, K, R$ with proper necessary conditions. We thus extend the neoclassical analysis by a dynamic adjustment rule which explains in a stable case how the firm reaches its profit maximizing state with time. Later we see that this extension allows the modeling of permanent growth in this framework too, that is, the system may not be stable.

A relation, which fulfills the above rules, is

$$x'(t) = \Psi_x\left(\frac{\partial\Pi}{\partial x}\right), \quad \Psi'_x\left(\frac{\partial\Pi}{\partial x}\right) > 0, \quad \Psi_x(0) = 0, \quad x = P, L, K, R. \tag{3}$$

The first order Taylor series approximation of function Ψ_x in the neighborhood of the optimum point $\partial\Pi/\partial x = 0$ is

$$\Psi_x(y) \approx \Psi_x(0) + \Psi'_x(0)(y - 0) = \Psi'_x(0)y.$$

With this approximation, we can write Eq. (3) as

$$x'(t) = \Psi'_x(0)\frac{\partial\Pi}{\partial x} \quad \Leftrightarrow \quad m_x x'(t) = \frac{\partial\Pi}{\partial x} \quad \text{where} \quad \Psi'_x(0) = \frac{1}{m_x}. \tag{4}$$

Now $L'(t)$ and $R'(t)$ with unit h/y^2 correspond to the acceleration of the use of both types of labor of the firm. Following Newton, we interpret the positive constants $m_x = \Psi'_x(0)$, $x = L, R$ as the inertial factors (the 'mass') of the adjusting quantities. These factors originate from legislation for overtime work, time needed to find skillful workers and employ them etc. The measurement unit $(eur \times y^2)/h^2$ of factors m_x, $x = L, R$ makes Eq. (4) dimensionally homogeneous (both sides have equal unit). With these assumptions, Eq. (4) with $x = L, R$ exactly corresponds to the Newtonian formulation for dynamics: $ma = F$, where $a = x'(t)$ and $F = \partial\Pi/\partial x$ with unit eur/h.

Eq. (4) with $x = P, K$ is not analogous to the Newtonian formulation because $P'(t)$ and $K'(t)$ are the velocities of P and K. However, we still identify the constants $\Psi'_x(0) = 1/m_x$, $x = P, K$ as the inverses of the inertial factors related to the adjustment of P and K, and $\partial\Pi/\partial x$, $x = P, K$ with units $unit/y$ and $1/y$, respectively — the causes for these velocities — as the 'forces the firm's managers direct upon these quantities'. These inertial factors originate from time and costs needed for price changes, to install new machines etc. The higher these factors the greater are m_x, $x = P, K$ with units $unit^2/eur$ and $1/eur$, respectively. Weiss (1993), for example, find empirical evidence of price inertia in concentrated industries.

The system (4) is then

$$\begin{pmatrix} m_P & 0 & 0 & 0 \\ 0 & m_L & 0 & 0 \\ 0 & 0 & m_R & 0 \\ 0 & 0 & 0 & m_K \end{pmatrix} \begin{pmatrix} P'(t) \\ L'(t) \\ R'(t) \\ K'(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial\Pi}{\partial P} \\ \frac{\partial\Pi}{\partial L} \\ \frac{\partial\Pi}{\partial R} \\ \frac{\partial\Pi}{\partial K} \end{pmatrix}. \tag{5}$$

In Estola (2001) we studied the role of increasing demand, decreasing costs with time and increasing returns to scale for the growth of production of a profit-seeking firm. For this reason, we skip the increasing demand case and assume in the following that the unit price of the produced good stays fixed, $P'(t) = 0$. The block diagram representation of this model is in Figure 1.

In the following we completely omit uncertainty to keep the modeling as simple as possible. However, it is relatively easy to add uncertainty into the analysis by replacing the quantities firm's managers consider in their decision-making by their expected values, see Estola (2001). In order to separate the reasons which cause decreasing costs with time and increasing returns to scale, we study every case by a separate model.

5 Increasing Efficiency due to Learning

Suppose a firm uses only labor input in its production. The costs of the firm then only consist of the working hours with wage w_1 (eur/h). The efficiency of labor — measured by the amount of goods produced in one hour with fixed labor input $S(t)$ $(unit/h)$ — is assumed to increase with time as

$$S(t) = bArcTan(t/t_0), \quad t \geq 0,$$

where b is a positive constant with unit $unit/h$, time t is measured in units y and $t_0 = 1$ (y). The argument of the transcendental function is thus dimensionless, as it should be. With t/t_0 from 0 to ∞, $ArcTan$ is increasing and takes values from 0 to 1.5 with a decreasing rate of growth. The shape of the function mimics a learning process where most learning takes place at the beginning and learning ends in a finite time. The maximum efficiency of labor is $1.5b$ $(unit/h)$ after all possible learning.

The production function of the firm is

$$q(t) = S(t)L(t)$$

and the profit function of the firm $\Pi(t)$ (eur/y) is

$$\Pi(t) = PS(t)L(t) - (1+\tau)w_1L(t),$$

where the unit price P and wage w_1 are assumed constant. According to Eq. (4), the following equation results for the dynamics of the labor input

$$m_L L'(t) = PbArcTan(t/t_0) - (1+\tau)w_1, \quad t \geq 0. \tag{6}$$

Now, engineering control systems base on Newton's dynamic laws of nature, and in the forces there exists parameters which the controller can adjust. Now if government knows the 'dynamic law of a firm' (6), it can operate with parameter τ as an open-loop controller of the process as in engineering. However, system (6) is not completely controllable because the exogenous factors P, w_1 together with time t affect the adjustment.

The solution of (6) is

$$L(t) = L(0) + \frac{Pbt}{m_L}ArcTan(t/t_0) - \frac{(1+\tau)w_1}{m_L}t - \frac{1}{2}\frac{Pbt_0Log(1+(t/t_0)^2)}{m_L},$$

where $L(0) = 0$. With proper values of P, w_1, b, permanent growth in $L(t)$ takes place with time and $q(t)$ expands according to $q(t) = S(t)L(t)$.

We can remark here that various other ways exist to add learning into the production process. Possible endogenous ways are, for example, to assume that the efficiency of labor $S(t)$ depends on the experience gained during the history of production, measured, for example, by the accumulated amount of production $\int_0^t q(s)ds$ or the accumulated working hours $\int_0^t L(s)ds$. These give similar results as above, but in these cases it is difficult to find an analytic solution for Eq. (6). All these three cases are shown in the block diagram representation of this model in Figure 2. For more accurate results concerning the real world, an increase in the wage level due to the increasing productivity of labor should be taken account.

6 Increasing Efficiency due to Capital

Next we show that growth may occur due to increasing returns to scale in the traditional inputs. We assume a Cobb-Douglas type of production function

$$q(t) = f(K(t), L(t)) = A\left(\frac{L(t)}{L_0}\right)^{\alpha}\left(\frac{K(t)}{K_0}\right)^{\beta},$$

where A with unit $unit/y$ is a technology constant, L_0, K_0 with units h/y and eur, respectively, are the amounts of the two inputs at time moment 0 and α, β are positive numbers. In this form the function is dimensionally homogeneous, and the marginal productivities of the two inputs are

$$\begin{aligned}\frac{\partial q}{\partial L} &= \frac{\alpha A}{L_0}\left(\frac{K(t)}{K_0}\right)^{\beta}\left(\frac{L(t)}{L_0}\right)^{\alpha-1}\left(\frac{unit}{h}\right),\\ \frac{\partial q}{\partial K} &= \frac{\beta A}{K_0}\left(\frac{L(t)}{L_0}\right)^{\alpha}\left(\frac{K(t)}{K_0}\right)^{\beta-1}\left(\frac{unit/y}{eur}\right).\end{aligned}$$

Factor $(K(t)/K_0)^{\beta}$ is thus a coefficient in the marginal productivity of labor, and factor $(L(t)/L_0)^{\alpha}$ a coefficient in the marginal productivity of capital.

The profit function of the firm measured in units eur/y is

$$\Pi(t) = PA\left(\frac{K(t)}{K_0}\right)^{\beta}\left(\frac{L(t)}{L_0}\right)^{\alpha} - (1+\tau)w_1 L(t) - rK(t).$$

The adjustment system for the inputs is

$$\begin{pmatrix} m_L & 0 \\ 0 & m_K \end{pmatrix}\begin{pmatrix} L'(t) \\ K'(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial \Pi}{\partial L} \\ \frac{\partial \Pi}{\partial K} \end{pmatrix}, \tag{8}$$

where the components of the force vector are:

$$\begin{aligned}\frac{\partial \Pi}{\partial L} &= \frac{\alpha PA}{L_0}\left(\frac{L(t)}{L_0}\right)^{\alpha-1}\left(\frac{K(t)}{K_0}\right)^{\beta} - (1+\tau)w_1,\\ \frac{\partial \Pi}{\partial K} &= \frac{\beta PA}{K_0}\left(\frac{L(t)}{L_0}\right)^{\alpha}\left(\frac{K(t)}{K_0}\right)^{\beta-1} - r.\end{aligned}$$

The force vector shows how the exogenous variables P, w_1, r affect the adjustment, and how government can affect the labor input by parameter τ. The block diagram of system (8) is in Figure 3 where capital is assumed to affect the production process via the productivity of labor. Due to the nonlinearity of system (8), its dynamic behavior is studied by phase diagrams.

The slopes of the demarcation lines $L'(t) = K'(t) = 0$ in coordinates (L, K) are

$$\begin{aligned} \left.\frac{dK}{dL}\right|_{L'(t)=0} &= -\frac{(\alpha-1)}{\beta}\frac{K}{L}, \\ \left.\frac{dK}{dL}\right|_{K'(t)=0} &= -\frac{\alpha}{(\beta-1)}\frac{K}{L}, \quad L, K > 0, \end{aligned}$$

and the attract/repel -character of the lines depend on the partials:

$$\begin{aligned} \frac{\partial L'(t)}{\partial L} &= \frac{\alpha(\alpha-1)PA}{L_0^2 m_L}\left(\frac{K(t)}{K_0}\right)^{\beta}\left(\frac{L(t)}{L_0}\right)^{\alpha-2}, \\ \frac{\partial K'(t)}{\partial K} &= \frac{\beta(\beta-1)PA}{K_0^2 m_K}\left(\frac{L(t)}{L_0}\right)^{\alpha}\left(\frac{K(t)}{K_0}\right)^{\beta-2}. \end{aligned}$$

The behavior of the system thus critically depends on the returns to scale of the two inputs, see the Appendix Part B.

Case 1: Decreasing Returns to Scale in Both Inputs

Decreasing returns to scale in both inputs corresponds to $0 < \alpha, \beta < 1$. Both demarcation lines are then upward sloping in coordinates (L, K) and the possible phase diagrams are in Figures 4, 5. The equilibrium $L'(t) = K'(t) = 0$ is stable in Figure 4 ($\alpha + \beta < 1$) and a saddle in Figure 5 ($\alpha + \beta > 1$). Figure 4 describes a dynamic extension for the neoclassical theory. Figure 5, on the other hand, shows that permanent growth in the use of the two inputs (and also in q through the production function) is possible if the initial values L_0, K_0 are great enough, while low values lead to the collapse of the firm. The profitability of a growing firm will increase with time.

Case 2: Increasing Returns to Scale in Capital

Decreasing returns to scale in labor and increasing returns to scale in capital correspond to $\alpha < 1, \beta > 1$. Now line $L'(t) = 0$ is upward and line $K'(t) = 0$ downward sloping in coordinates (L, K). The phase diagram in Figure 6 shows that the equilibrium is a saddle. Growth is thus possible if L_0, K_0 are great enough and K_0 is more important in this. However, low initial values lead to the collapse of the firm. Only firms with K_0 great enough can thus survive and their profitability increases with time.

Case 3: Increasing Returns to Scale in Labor

Decreasing returns to scale in capital and increasing returns to scale in labor correspond to $\alpha > 1, \beta < 1$. Line $L'(t) = 0$ is then downward and line $K'(t) = 0$ upward sloping in coordinates (L, K). The phase diagram in

Figure 7 shows that the equilibrium is a saddle. Growth is thus possible if L_0 is great enough while that of K_0 is not that important. However, low value of L_0 leads to the collapse of the firm. Only firms with L_0 great enough can thus survive and their profitability increases with time.

Case 4: Increasing Returns to Scale in Both Inputs

Increasing returns to scale in capital and labor correspond to $\alpha, \beta > 1$. Both demarcation lines are then downward sloping in coordinates (L, K). The only possible phase diagram is in Figure 8 because the case $L'(t) = 0$ steeper corresponds to $\alpha + \beta < 1$, which is impossible. The equilibrium in Figure 8 is a saddle. Permanent growth is possible if L_0, K_0 are great enough while low initial values lead to the collapse of the firm. Firms with L_0, K_0 great enough can thus survive and their profitability increases with time.

In the phase diagrams, changes in the exogenous variables P, w_1, r, τ move the positions of the demarcation lines. Thus if the initial state of the system is equilibrium, a move in a demarcation line starts an adjustment process toward a new equilibrium or away from it depending on the values of α, β.

7 Increasing Efficiency by Human Capital

Next we omit physical capital and concentrate on the role of human capital in the production process. The production function of the firm is

$$q(t) = BH(t)\left(\frac{L(t)}{L_0}\right)^{\alpha}, \quad H(t) = D\left(\frac{R(t)}{R_0}\right)^{\gamma},$$

where L (h/y) is the labor input in the ordinary production, R (h/y) the labor input in the production of human capital H (eur/y) and B $(unit/eur)$, D (eur/y), α, γ are positive constants. Notice that H is the velocity, and not the level, of human capital, and we assume that only increases in human capital affect q. Human capital is measured by its monetary value when sold to other firms. The profit function is

$$\Pi(t) = PA\left(\frac{R(t)}{R_0}\right)^{\gamma}\left(\frac{L(t)}{L_0}\right)^{\alpha} - (1+\tau)[w_1L(t) + w_2R(t)], \quad A = BD.$$

The model is thus identical as in the previous section if R is replaced by K, γ by β and $(1+\tau)w_2$ by r. Now either of the two labor types may obey increasing returns to scale, or they may together cause increasing returns to scale in the production. We could have included in the analysis physical capital too, and get results where every combination of the three inputs may cause increasing returns in the production. Various other forms for the production function could also have been assumed with similar results.

The endogenous growth theories of Lucas (1988) and Romer (1990) completely base on increasing returns to scale in production due to human capital. However, our modeling shows that the somewhat unclear concept human capital is not a necessity for economic growth. Similar results are obtained with increasing productivity of labor due to learning by doing at work, or because labor and capital together obey increasing returns to scale. The block diagram representations in Figures 2 and 3 give a hint of how a more detailed description of the production process could reveal elements of it where efficiency can be increased. However, in the next section we present a slightly different kind of a model where capital contributes in the efficiency of labor by reducing the required work time.

8 Increasing Efficiency in Working Time

Suppose that the time needed to produce one unit of the product $Z(t)$ $(h/unit)$ depends on the capital $K(t)$ (eur) workers have available:

$$Z(t) = z_0 + z_1 e^{-aK(t)},$$

where z_0, z_1, a are positive constants with units $h/unit$, $h/unit$ and $1/eur$, respectively. The required working time $L(t)$ (h/y) for the velocity of production q $(unit/y)$ is then

$$L(t) = Z(t)q(t) = q(t)[z_0 + z_1 e^{-aK(t)}].$$

With wage w_1 (eur/h), the labor costs of the firm during time unit y are

$$(1+\tau)w_1 L(t) = (1+\tau)w_1 q(t)[z_0 + z_1 e^{-aK(t)}].$$

The profit function of the firm is then

$$\begin{aligned}\Pi(t) &= Pq(t) - (1+\tau)w_1 L(t) - rK(t) \\ &= \Big[P - (1+\tau)w_1[z_0 + z_1 e^{-aK(t)}]\Big]q(t) - rK(t). \qquad (10)\end{aligned}$$

In last form of (10) we eliminated the labor input L, but now the managers of the firm can affect the profit by adjusting q and K. The profit-seeking adjustments as in (4) are

$$\begin{pmatrix} m_q & 0 \\ 0 & m_K \end{pmatrix} \begin{pmatrix} q'(t) \\ K'(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial \Pi}{\partial q} \\ \frac{\partial \Pi}{\partial K} \end{pmatrix}. \qquad (11)$$

The force which causes the acceleration of production $q'(t)$ is $\partial\Pi/\partial q$ ($eur/unit$) and m_q ($(eur \times y^2)/unit^2$) is the corresponding inertial factor. Due to the nonlinearity of system (11), its behavior is analyzed by a phase diagram. Figure 9 shows that the equilibrium is a saddle; see the Appendix Part C. Permanent growth will thus occur if q_0, K_0 are great enough, that is, the firm has enough capital to reduce the working time required at such level that producing one good is profitable, and the velocity of production is great enough for the revenues to cover the interest costs of the capital. With low values of q_0, K_0, the firm will collapse.

9 Conclusions

Static neoclassical theory of a firm, where time is eliminated and which models only the profit maximizing state, does not allow increasing efficiency with time or increasing returns to scale. A different framework is thus needed for the modeling of time dependent economic processes. As a candidate we introduced a dynamic extension for the static neoclassical theory, analogous to the Newtonian formulation in physics. The constructed models are of control theoretic nature which we demonstrated by their block diagrams. The main results from the models are: 1) Permanent growth may occur due to increasing efficiency with time or increasing returns to scale in any input or a combination of inputs of a firm. 2) The input mix of a growing firm will change with time toward greater profitability. 3) Government may act with taxes as an open-loop controller of firms' production processes.

Naturally, there exist factors which slow down and eventually stop the growth of a firm, like increasing wage level, the saturation of a market, the appearance of competing firms etc. However, we have evidence of long-lasting growth of business firms such as IBM, HP, Nokia etc. Even though these growth processes may not last forever, we believe that it is meaningful to model these as explosive processes rather than assume that the firms have been approaching their equilibrium states last 20 years, as the neoclassical theory claims. We believe that there are elements in the production processes of these firms which allow them to grow in a profitable way, which elements we tried to model here. However, because firms' production methods change with time, one specific model may not keep its accuracy very long.

From the point of view of economic policy, we showed that public sector can affect firms' use of inputs by wage tax. Other tax forms and subsidies could have been assumed but were omitted to avoid complexity. A decrease in interest rate, if controllable by monetary authorities, can also be used to promote growth, as the defined forces show.

Appendix

Part A: The continuous time interest rate r $(1/y)$ for time unit y is defined as the growth rate of a monetary quantity $x(t)$ (eur)

$$r = \frac{x'(t)}{x(t)},$$

where the unit of $x'(t)$ is eur/y. This explains the measurement unit of r. Assuming r fixed, the solution of this differential equation is

$$x(t) = x(t_0)e^{r\times(t-t_0)}.$$

Setting $t_0 = 0$, the above interest factor becomes e^{rt} where time t is measured in units y. The exponent of e is thus dimensionless as it should be.

Part B: The necessary conditions for the critical point $\frac{\partial\Pi}{\partial L} = \frac{\partial\Pi}{\partial K} = 0$ to maximize the profit are:

$$\begin{aligned}
\frac{\partial\Pi}{\partial L} &= \frac{PA\alpha}{L_0}\left(\frac{L}{L_0}\right)^{\alpha-1}\left(\frac{K}{K_0}\right)^{\beta} - (1+\tau)w_1 = 0,\\
\frac{\partial\Pi}{\partial K} &= \frac{PA\beta}{K_0}\left(\frac{L}{L_0}\right)^{\alpha}\left(\frac{K}{K_0}\right)^{\beta-1} - r = 0.
\end{aligned}$$

The sufficient condition for profit maximization is that the first minor determinant of the Hessian matrix is negative and the second is positive. The Hessian matrix is

$$H = \begin{pmatrix} \frac{\partial^2\Pi}{\partial L^2} & \frac{\partial^2\Pi}{\partial L\partial K} \\ \frac{\partial^2\Pi}{\partial K\partial L} & \frac{\partial^2\Pi}{\partial K^2} \end{pmatrix}.$$

Its first minor is $H_1 = \frac{\partial^2\Pi}{\partial L^2}$ and the second minor equals the determinant of the whole Hessian,

$$\begin{aligned}
|H| &= \frac{\partial^2\Pi}{\partial L^2}\frac{\partial^2\Pi}{\partial K^2} - \left(\frac{\partial^2\Pi}{\partial L\partial K}\right)^2, \quad \text{where}\\
\frac{\partial^2\Pi}{\partial L^2} &= \frac{PA\alpha(\alpha-1)}{{L_0}^2}\left(\frac{K}{K_0}\right)^{\beta}\left(\frac{L}{L_0}\right)^{\alpha-2},\\
\frac{\partial^2\Pi}{\partial K^2} &= \frac{PA\beta(\beta-1)}{{K_0}^2}\left(\frac{L}{L_0}\right)^{\alpha}\left(\frac{K}{K_0}\right)^{\beta-2},\\
\frac{\partial^2\Pi}{\partial L\partial K} &= \frac{PA\alpha\beta}{L_0K_0}\left(\frac{L}{L_0}\right)^{\alpha-1}\left(\frac{K}{K_0}\right)^{\beta-1}.
\end{aligned}$$

Thus

$$\begin{aligned}
|H| &= \frac{P^2A^2\alpha\beta(\alpha-1)(\beta-1)}{{L_0}^2{K_0}^2}\left(\frac{L}{L_0}\right)^{2(\alpha-1)}\left(\frac{K}{K_0}\right)^{2(\beta-1)} \\
&- \frac{P^2A^2\alpha^2\beta^2}{{L_0}^2{K_0}^2}\left(\frac{L}{L_0}\right)^{2(\alpha-1)}\left(\frac{K}{K_0}\right)^{2(\beta-1)} \\
&= \frac{P^2A^2\alpha\beta}{{L_0}^2{K_0}^2}\left(\frac{L}{L_0}\right)^{2(\alpha-1)}\left(\frac{K}{K_0}\right)^{2(\beta-1)}[(\alpha-1)(\beta-1)-\alpha\beta].
\end{aligned}$$

Now if $0 < \alpha < 1$, then $H_1 < 0$. On the other hand, requirement $|H| > 0$ corresponds to $\alpha+\beta < 1$, and thus for the critical point to be a maximum — as is assumed in the neoclassical framework — neither of the inputs separately or together can have increasing returns to scale. The stability of the system, too, depends on the Hessian matrix, see for example McCafferty (1990 p. 138). The trace of the Hessian matrix is

$$Tr(H) = \frac{\partial^2\Pi}{\partial L^2} + \frac{\partial^2\Pi}{\partial K^2} = PA\left(\frac{L(t)}{L_0}\right)^{\alpha}\left(\frac{K(t)}{K_0}\right)^{\beta}\left(\frac{\alpha(\alpha-1)}{L^2} + \frac{\beta(\beta-1)}{K^2}\right).$$

The system is stable if $|H| > 0 \Leftrightarrow (\alpha-1)(\beta-1) > \alpha\beta$ i.e. $\alpha+\beta < 1$ and $Tr(H) < 0$, which occurs certainly when $0 < \alpha, \beta < 1$; see Figure 4.

Part C: The partial derivatives of the profit function are:

$$\begin{aligned}
\frac{\partial\Pi}{\partial q} &= P - (1+\tau)w_1[z_0 + z_1e^{-aK}] = 0, \\
\frac{\partial\Pi}{\partial K} &= (1+\tau)w_1z_1aqe^{-aK} - r = 0.
\end{aligned}$$

The slopes of the demarcation lines are:

$$\begin{aligned}
\left.\frac{\partial K}{\partial q}\right|_{q'(t)=0} &= 0, \\
\left.\frac{\partial K}{\partial q}\right|_{K'(t)=0} &= \frac{(1+\tau)w_1z_1}{re^{aK}} > 0.
\end{aligned}$$

The attract/repel -character of the lines depend on the partials

$$\begin{aligned}
\frac{\partial q'(t)}{\partial K} &= \frac{(1+\tau)w_1z_1ae^{-aK}}{m_q} > 0, \\
\frac{\partial K'(t)}{\partial K} &= -\frac{(1+\tau)w_1z_1a^2qe^{-aK} - r}{m_K} < 0.
\end{aligned}$$

10 References

Arrow, K. J. & Hahn, F. H. (1971). General Competitive Analysis. San Francisco, California Holden-Day, Inc.

Bensoussan, A., Hurst E. G. Jr., Näslund, B. (1974). Management Applications of Modern Control Theory. North-Holland Publishing Company, Amsterdam.

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

de Jong, F. (1967). Dimensional Analysis for Economists. Amsterdam, North-Holland.

Dixit, A. K. & Pindyck, R. S. (1994). Investment Under Uncertainty. New Jersey, Princeton University Press.

Estola, M. (2001). A Dynamic Theory of a Firm: An Application of 'Economic Forces'. Advances in Complex Systems, Vol. 1, N:o 1, 163-176.

Fisher, F. M. (1983). Disequilibrium Foundations of equilibrium Economics. Cambridge University Press.

Lucas, R. E. Jr. (1988). On the Mechanics of Economic Development. Journal of Monetary Economics, 22, 3–42.

McCafferty, S. (1990). Macroeconomic Theory. Harper & Row, Publishers, New York.

Marshall, S. A. (1978). Introduction to Control Theory. MacMillan Press LTD.

Mas-Colell, A., M. D. Whinston and J. R. Green (1995). Microeconomic Theory. New York, Oxford University Press.

Mirowski, P. (1989a). The Rise and Fall of the Concept of Equilibrium in Economic Analysis. Recherches Economiques du Louvain, Vol. 55:4.

Mirowski, P. (1989b). More Heat than Light, Economics as Social Physics, Physics as Nature's Economics. Cambridge University Press.

Ogata, K. (1997). Modern Control Engineering, Third Edition. Prentice-Hall, International (UK) Limited, London.

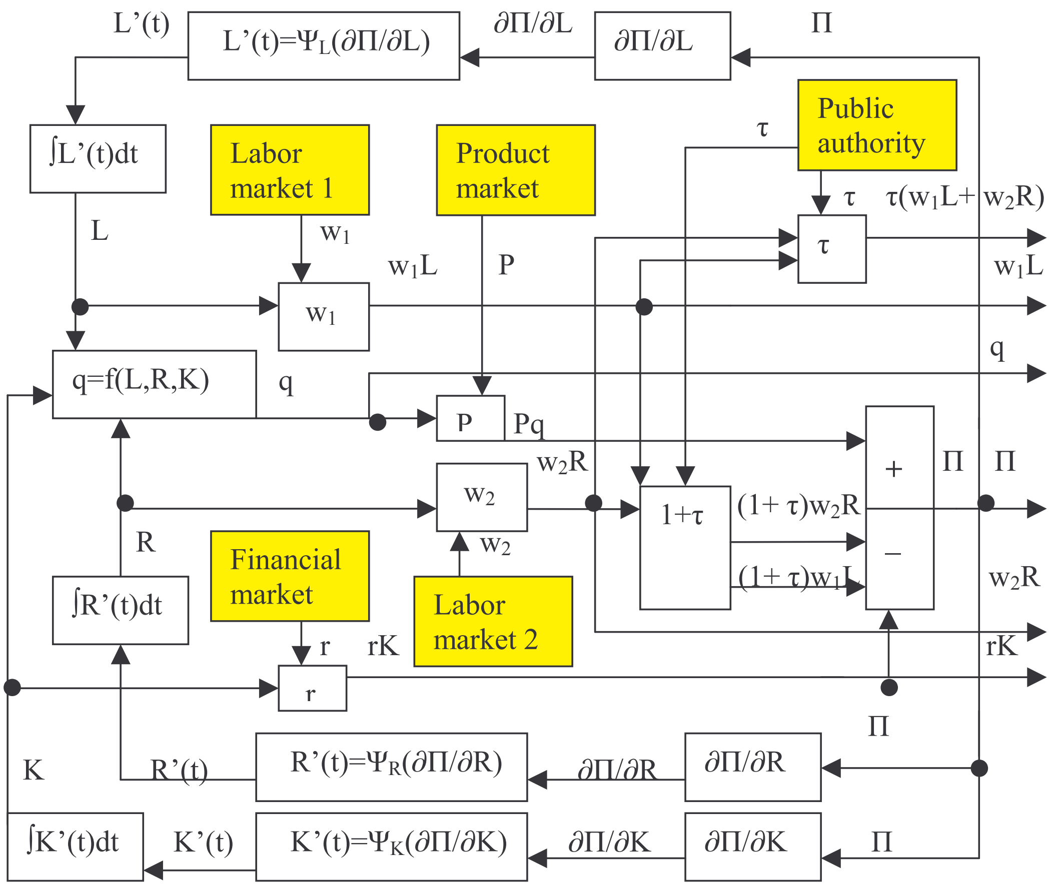
Romer, P. M. (1990). Endogenous Technological Change. Journal of Political Economy, October 98, 71-102.

Samuelson, P. A. (1942). The stability of equilibrium. Econometrica Vol. 10, 1–25. In book: The Collected Scientific Papers of Paul A. Samuelson, Volume I, ed. by Joseph E. Stiglitz, (1966). Cambridge Mass, M.I.T. Press.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. Quarterly Journal of Economics, Vol. 70, 65-94.

Weiss, C. (1993). Price inertia and market structure: empirical evidence from Austrian manufacturing. Applied Economics, Vol. 25, 1175–1186.

Figure 1. Block Diagram of the Production System of a Firm

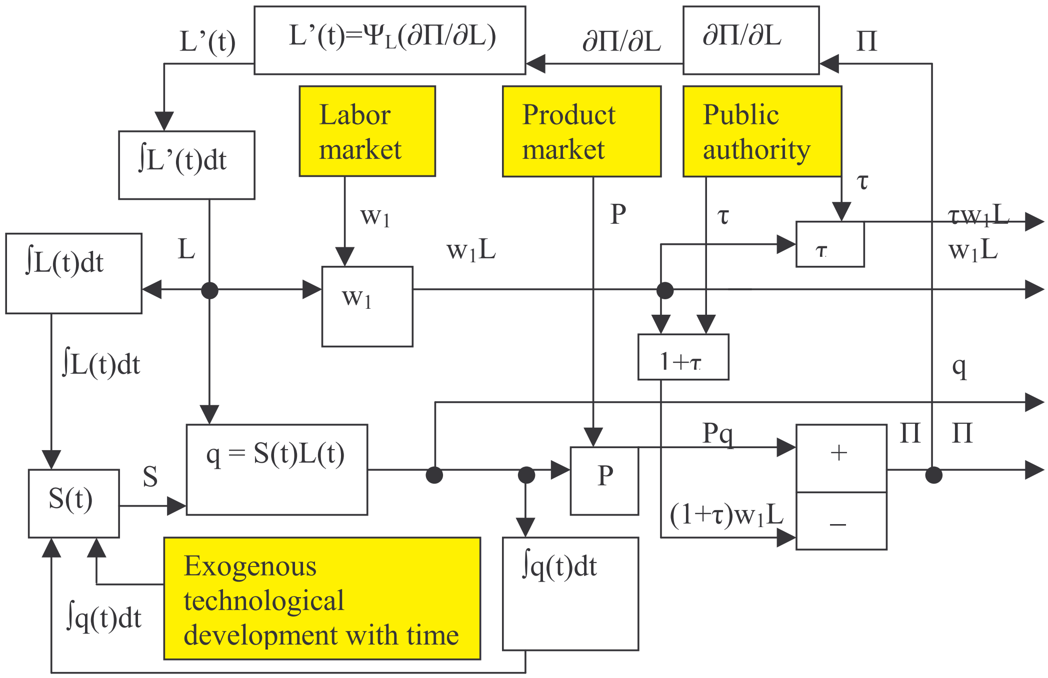


In the figure the 'flows' denoted by arrows are named by the symbols of the 'flowing' quantities. Mark ● means that the crossing 'flows' are connected, that is, the same variable is in both lines; otherwise crossing lines are assumed not to interact each other. White rectangles affect the 'flows' as is shown, and the grey rectangles are sources of external impulses into the system. For every rectangle the input and output flows are named. In the single [+/-] –box, the monetary flow entering the upper part is added with a positive sign, those entering the lower part are added with a negative sign. The output flow from this box is the profit Π. The diagram shows how profit-making, the ultimate goal of the firm, controls the production process. All three inputs in the production function are adjusted by a closed-loop mechanism, where the feed-back is based on the marginal profitability of the input. The real output of the system is q, and the monetary outputs are w_1L and w_2R for the two types of labor, $\tau(w_1L+w_2R)$ for the government, rK for the lenders of foreign capital or investors of own capital and Π to the firm's owners. The main definitions are:

$q = f(L,R,K)$, $\Pi = Pq - (1+\tau)w_1L - (1+\tau)w_2R - rK$, $\partial\Pi/\partial L = P(\partial f/\partial L) - (1+\tau)w_1$,

$\partial\Pi/\partial R = P(\partial f/\partial R) - (1+\tau)w_2$, $\partial\Pi/\partial K = P(\partial f/\partial K) - r$, $x'(t) = \Psi_x(\partial\Pi/\partial x)$, $x = L, R, K$.

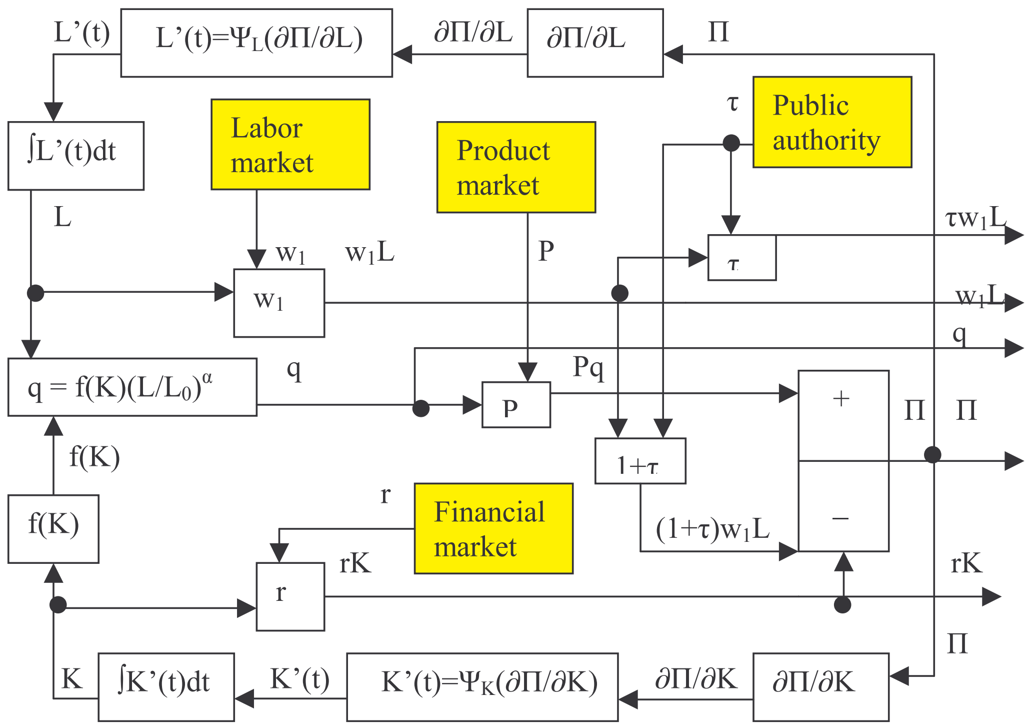
Figure 2. Block Diagram of a Production System with Learning



See the explanations in Figure 1. In Figure 2, technological development may occur due to exogenous (the grey box) or endogenous reasons (the white boxes representing the accumulated amounts of production and working hours). Endogenous technological development takes place via closed-loop feed-back mechanisms, and exogenous technological development through an open-loop mechanism. Labor is adjusted according to the marginal profitability by a closed-loop feed-back mechanism, and the markets for the unit price and wage open the system for external impulses. The main definitions are:

$q(t) = S(t)L(t)$, $\Pi(t) = P(t)q(t) - (1+\tau)w_1(t)L(t)$, $\partial\Pi/\partial L = P(t)S(t) - (1+\tau)w_1(t)$, $L'(t) = \Psi_L(\partial\Pi/\partial L)$.

Figure 3. Block Diagram of a Production System with L and K



See the explanations in Figure 1. The main definitions are given below.

$q = f(K)(L/L_0)^{\alpha}$, $\Pi = Pq - (1+\tau)w_1L - rK$, $\partial\Pi/\partial L = \alpha Pf(K)(L/L_0)^{\alpha-1}/L_0 - (1+\tau)w_1$,

$\partial\Pi/\partial K = P(L/L_0)^{\alpha}f'(K) - r$, $x'(t) = \Psi_x(\partial\Pi/\partial x)$, $x = L, K$.

Figure 4

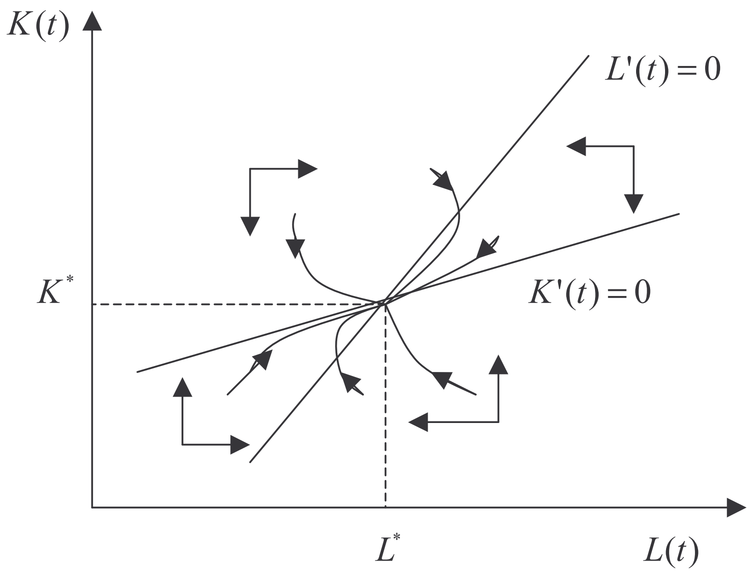


Figure 5

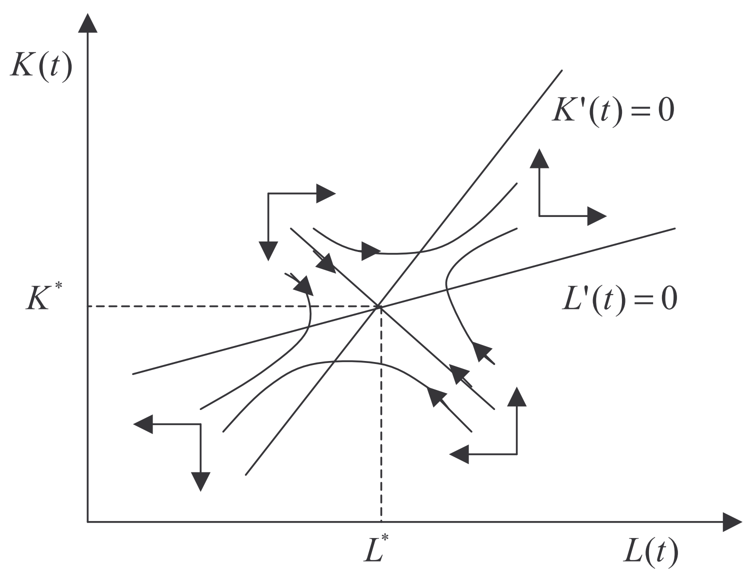


Figure 6

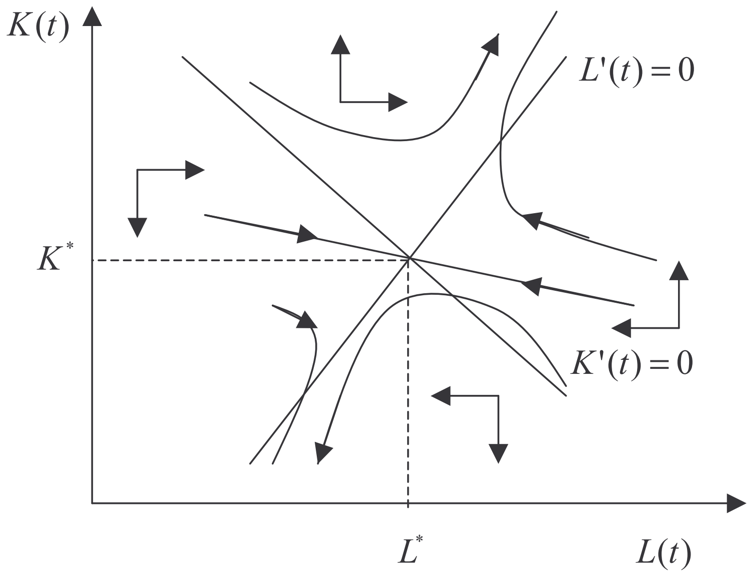


Figure 7

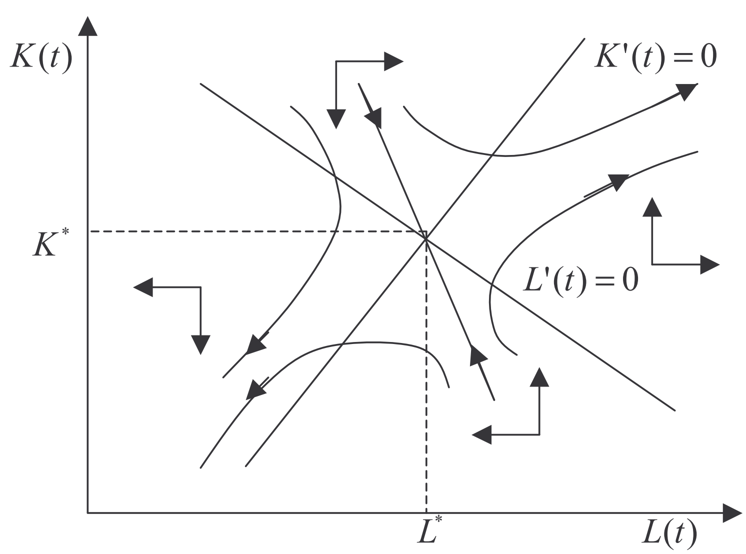


Figure 8

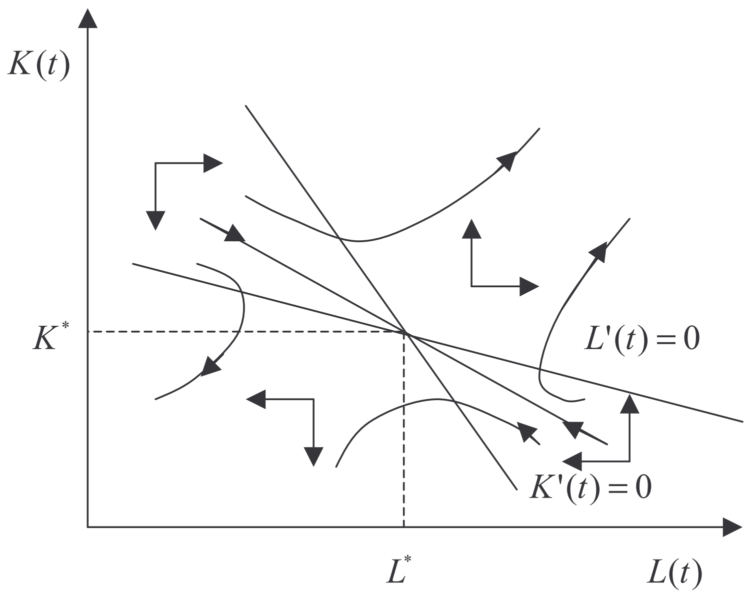


Figure 9