9. Externalities, Human Capital and Endogenous Growth
Economic Growth and Learning by Doing

We now turn to a growth model which is based on the assumption of positive externalities from aggregate capital accumulation on labor efficiency. The main idea that drives this model is learning by doing, an idea introduced to growth models by Arrow (1962). This assumption can, under certain conditions, lead to endogenous growth, as in Romer (1986).

In the learning by doing model, labor efficiency is a function of both exogenous technical progress, as well as aggregate capital per worker. Thus, the efficiency of labor, which is the same for all firms, depends on capital per worker in the rest of the economy. Because of learning by doing, as suggested by Arrow (1962), the accumulation of aggregate capital increases labor productivity both directly and indirectly, through “knowledge spillovers”, that have a direct effect on the efficiency of labor. It is assumed here that "knowledge" is like a public good, and that the accumulation of knowledge depends on the accumulation of aggregate capital.

An important consequence of this approach is that diminishing returns from capital accumulation set in more slowly, and that, under certain conditions, there may even be constant or increasing returns from capital accumulation. In these latter circumstances growth becomes endogenous and is determined by the rate of accumulation of aggregate physical capital.
Learning by Doing and Capital Accumulation

“I advance the hypothesis here that technical change in general can be ascribed to experience, that it is the very activity of production which gives rise to problems for which favorable responses are selected over time. … I therefore take … cumulative gross investment (cumulative production of capital goods) as an index of experience. Each new machine produced and put into use is capable of changing the environment in which production takes place, so that learning is taking place with continually new stimuli. This at least makes plausible the possibility of continued learning in the sense, here, of a steady rate of growth in productivity.” Arrow (1962).
The Production Function of Individual Firms

Production of goods and services takes place through a large number of competitive firms. The production function of firm $i$ is given by,

$$Y_i(t) = AF(K_i(t), h(t)L_i(t)), i = 1, 2, ...$$

The production function is characterized by constant returns to scale and satisfies the usual conditions of a neoclassical production function. We can thus write the production function as,

$$\hat{y}_i(t) = Af(\hat{k}_i(t), h(t))$$

where, $\hat{y}_i = Y_i/L_i$, output per worker of firm $i$, $\hat{k}_i = K_i/L_i$, physical capital per worker of firm $i$, $h$, efficiency of labor, or human capital per worker in firm $i$, assumed to be the same for all firms.

In what follows we shall assume a Cobb-Douglas production function of the form,

$$Y_i(t) = AK_i(t)^\alpha(h(t)L_i(t))^{1-\alpha}$$

where, $0 < \alpha < 1$ is the direct elasticity of output with respect to the capital stock, and $A$ is total factor productivity. Those two parameters, as well as the efficiency of labor $h$, are assumed to be the same for all firms.
Externalities from the Accumulation of Capital

The efficiency of labor $h$ is determined by,

$$h(t) = \left( \frac{K(t)}{L(t)} \right)^\beta \left( e^{gt} \right)^{1-\beta}$$

where $0 \leq \beta \leq 1$. $K$ is the aggregate capital stock (an index of experience) and $L$ is aggregate employment of labor. A higher aggregate capital stock per worker implies a higher efficiency of labor for all workers in the economy, irrespective of the capital stock employed by the firm that employs the particular worker. The parameter $\beta$ measures the elasticity of labor efficiency with respect to the aggregate capital stock per worker, and $g$ is the exogenous rate of technical progress.

Substituting for the efficiency of labor in the production function and aggregating across firms, we get,

$$Y(t) = A(K(t))^{\alpha+\beta(1-\alpha)}(e^{gt}L(t))^{1-(\alpha+\beta(1-\alpha))}$$

It follows that aggregate output per worker is given by,

$$\hat{y}(t) = A(\hat{k}(t))^{\alpha+\beta(1-\alpha)}(e^{gt})^{1-\alpha-\beta(1-\alpha)}$$

where $\hat{y} = Y/L$ and $\hat{k} = K/L$. 

George Alogoskoufis, Dynamic Macroeconomics, 2019
Output per Worker as a Function of Capital per Worker

\[ y = \beta k \]

- \( \beta > 1 \)
- \( \beta = 1 \)
- \( 0 < \beta < 1 \)
- \( \beta = 0 \)
Externalities from the Accumulation of Capital and the Aggregate Production Function

For $\beta=0$, there is no effect of the aggregate accumulation of capital on labor efficiency, and we are back to an aggregate Cobb Douglas production function without externalities. The efficiency of labor depends only on the exogenous rate of technical progress $g$.

For $0<\beta<1$ the accumulation of capital implies a positive externality on the efficiency of labor, but the aggregate marginal product of capital tends to fall as the economy accumulates more capital per worker. Capital accumulation leads to diminishing returns, although the productivity of capital declines at a slower rate than if there were no externalities.

For $\beta=1$, the aggregate marginal product of capital is constant, equal to total factor productivity $A$, and is not affected by the accumulation of capital. There are no diminishing returns to capital accumulation as the aggregate marginal product of capital is constant.

Finally, for $\beta>1$, the marginal product of capital increases with capital accumulation, but this assumption violates the condition of constant returns to scale for the aggregate economy.
We shall concentrate on the special case $\beta = 1$, which implies endogenous growth without violating the assumption of constant returns to scale.

In the special case $\beta = 1$, aggregate output and output per worker take the form,

$$Y(t) = AK(t)$$
$$\hat{y}(t) = \hat{k}(t)$$

Aggregate output per worker is a linear function of capital per worker, although this is not the case for individual firms. The marginal and average productivity of capital is constant. Because of the linearity of the aggregate production function, the rate of growth of output per worker, or per capita income $g$ is equal to the rate of growth of capital per worker. The accumulation of capital per worker does not lead to diminishing returns. It thus follows that,

$$\frac{\hat{y}(t)}{\hat{y}(t)} = \frac{\hat{k}(t)}{\hat{k}(t)} = g$$

where $g$ is now endogenous, and is determined by the rate of accumulation of capital.

The rate of growth of aggregate output $g + n$, is determined by the share of net investment to total output,

$$g + n = \frac{\dot{Y}(t)}{Y(t)} = \frac{\dot{K}(t)}{K(t)} = \frac{(I(t)/K(t))}{(I(t)/K(t))} - \delta = A(I(t)/Y(t)) - \delta$$

where $I$ denotes gross investment in physical capital and $\delta$ the depreciation rate.
AK Models of Endogenous Growth

This endogenous growth model belongs to a class of models known as *AK Models*, from the form of the aggregate production function. The *AK* model with learning by doing was analyzed by Romer (1986), and we shall refer to it as the *Arrow-Romer model*.

An alternative *AK* model which is widely used is due to Rebelo (1991), who demonstrates that as long as there is a “core” of capital goods whose production does not involve non-reproducible factors, endogenous growth is compatible with production technologies that exhibit constant returns to scale. The simplest Rebelo (1991) model is a two sector model in which consumption goods are produced using both capital and labor, but capital goods are produced using only capital. In this class of models there are no externalities through “knowledge spillovers”, but in competitive equilibrium aggregate output turns out to be proportional to the capital stock. Because the Rebelo model is not based on externalities its policy implications are different than the policy implications of the Arrow-Romer model.
Determination of the Real Interest Rate and Real Wages

Firms operate under perfect competition and they maximize profits by taking the prices of inputs as given. Profit maximization implies that the real interest rate will be equal to the marginal product of capital for individual firms, and the real wage to the marginal product of labor for individual firms. These are given by,

\[
 r(t) = \alpha A \hat{k}_i(t)^{a-1} h(t)^{1-a} - \delta \\
 \hat{w}(t) = (1 - \alpha) A \hat{k}_i(t)^{a} h(t)^{1-a}
\]

The marginal productivity condition for the real interest rate implies that all firms will select the same capital per worker, as they face the same real interest rate and the same labor efficiency per worker. Therefore, \( \hat{k}_i(t) = \hat{k}(t), \forall i \).

Under the assumption that \( \beta = 1 \) we have that \( h(t) = \hat{k}(t) \),

As a result, the real interest rate will be equal to,

\[
 r(t) = \alpha A - \delta
\]

The real interest rate is constant and equal to the private marginal product of capital, as calculated by each individual firm. Because of its small size, each competitive firm does not internalize the effects of its own choice of capital per worker on the aggregate capital stock per worker in the economy, treating labor efficiency as exogenously given. Labor efficiency depends on the aggregate capital stock per worker and not on the firm specific capital per worker. Because of this externality each firm underestimates the social marginal product of capital, and the real interest rate, as determined in competitive markets, is lower than the social (aggregate) marginal product of capital.

The real wage per worker is given by,

\[
 \hat{w}(t) = (1 - \alpha) A \hat{k}(t) = (1 - \alpha) \hat{y}(t)
\]

The real wage is a constant share of output per worker. If output per worker is growing at a rate \( g \), then the real wage per worker will also be growing at a rate \( g \). These properties are in accordance with Kaldor’s stylised facts.
Endogenous Growth and the Savings Rate in the Learning by Doing Model

Let us first assume that, as in the Solow model, consumer behavior is described by a constant savings rate. Per capita consumption is thus given by,

\[ \hat{c}(t) = (1 - s)\hat{y}(t) \]

where \( s \) is the exogenous savings rate.

Per capita output is given by,

\[ \hat{y}(t) = A\hat{k}(t) \]

In equilibrium, total output will be equal to consumption plus gross investment. In per capita terms,

\[ \hat{y}(t) = \hat{c}(t) + \hat{k}(t) + (n + \delta)\hat{k}(t) \]

Substituting the consumption function and the production function in the equilibrium condition, it follows that,

\[ \hat{k}(t) = (sA - n - \delta)\hat{k}(t) \]

Dividing both sides \( \hat{k}(t) \) we get that,

\[ g = \frac{\hat{k}(t)}{\hat{k}(t)} = sA - (n + \delta) \]

This determines the endogenous growth rate \( g \). The higher the exogenous savings (and investment) rate \( s \), and total factor productivity \( A \), the higher the growth rate of per capita income. On the other hand, population growth and the depreciation rate have a negative impact on the endogenous growth rate.
Endogenous Growth and the Savings Rate in the Learning by Doing Model

In this endogenous growth model the savings rate plays a similar role to its role in the exogenous growth Solow model. In the exogenous growth Solow model, the savings rate has a positive impact on steady state capital per effective unit of labor $k^*$, and the growth rate during the convergence process towards $k^*$, but does not affect the long-term growth rate, which is equal to the exogenous rate of technical progress $g$.

In this endogenous growth Solow model, with positive externalities from capital accumulation, the saving rate $s$ determines the gross investment rate, and, through the gross investment rate the endogenous growth rate.

The accumulation of capital does not imply diminishing returns for the marginal product of capital in this endogenous growth model, and growth continuous for ever.

Therefore, in this endogenous growth model, economies characterized by different stocks of initial capital, will not converge to the same per capita income, even if they have the same savings rate, the same total factor productivity, the same rate of population growth and the same depreciation rate. They will simply have the same steady state growth rate, without converging to the same per capita income on the balanced growth path. Their initial differences will be maintained for ever.
Endogenous Growth in a Representative Household Learning by Doing Model

From the Euler equation for consumption, which describes the optimal savings behavior of the representative household, the rate of change of per capita consumption is given by,

\[ \frac{\dot{c}(t)}{\bar{c}(t)} = \frac{1}{\theta} \left( r(t) - \rho \right) \]

Given that in this learning-by-doing endogenous growth model the real interest rate is constant and equal to \( \alpha A - \delta \), the rate of growth of per capita consumption is also constant, and given by,

\[ \frac{\dot{c}(t)}{\bar{c}(t)} = \frac{1}{\theta} \left( \alpha A - \delta - \rho \right) = g \]

This determines the endogenous growth rate \( g \), as on a balanced growth path all per capita variables grow at the same rate. If that was not so, either consumption or investment would eventually grow to be equal to output.

A higher pure rate of time preference of households, relative to the private marginal product of capital to firms (the real interest rate), results in a lower endogenous growth rate of per capita income and consumption. This is because a higher pure rate of time preference of households implies lower savings and a lower rate of accumulation of capital. On the other hand, a higher total factor productivity \( A \), or a higher private elasticity of firm output to capital \( \alpha \), lead to a higher endogenous growth rate, as both result in a higher equilibrium real interest rate and higher savings and capital accumulation rates. For the opposite reason, the depreciation rate \( \delta \) has a negative impact on the endogenous growth rate.

It is straightforward to show that in this model the savings rate is constant and determined by,

\[ s = 1 - \left( \frac{\dot{c}(t)}{\bar{y}(t)} \right) = \frac{n + g + \delta}{A} = \frac{1}{\theta A} \left( \alpha A - (1 - \theta)\delta - \rho + \theta n \right) \]
The Determinants of Endogenous Growth in the Representative Household Learning by Doing Model

A higher pure rate of time preference of households $\rho$, relative to the private marginal product of capital to firms (the real interest rate), results in a lower endogenous growth rate of per capita income and consumption. This is because a higher pure rate of time preference of households implies lower savings and a lower rate of accumulation of capital.

On the other hand, a higher total factor productivity $A$, or a higher private contribution of capital to output $\alpha$, lead to a higher endogenous growth rate, as both result in a higher equilibrium real interest rate, and higher savings and capital accumulation rates.

For the opposite reason, the depreciation rate $\delta$ has a negative impact on the endogenous growth rate.

A higher elasticity of inter-temporal substitution of consumption $1/\theta$ results in a higher growth rate, as it facilitates savings.
The Inefficiency of Competitive Equilibrium in the Endogenous Growth Ramsey Model

Let us assume there is a social planner who maximizes the intertemporal utility function of the representative household, under the aggregate and not the private capital accumulation constraint. The first order condition for an optimum would we given by,

$$\frac{\dot{c}(t)}{\hat{c}(t)} = \frac{1}{\theta} \left( A - \delta - \rho \right) = g^* > \frac{1}{\theta} \left( \alpha A - \delta - \rho \right) = g$$

From the above inequality we can deduce that the endogenous growth rate in the competitive economy is lower than the socially efficient growth rate. This is because the competitive real interest rate ($\alpha A - \delta$) underestimates the social net marginal product of capital ($A - \delta$), which, because of the positive externality from capital accumulation, is higher than the private net marginal product of capital. Since the positive externality from capital accumulation is not reflected in the real interest rate, households have a smaller incentive to save and accumulate capital, and, as a result, the savings rate, the investment rate and the growth rate of the economy are lower than what would be socially optimal.

Thus, the competitive equilibrium is sub-optimal, and could be improved upon by a social planner who would choose savings and capital accumulation taking the externality into account. Various other policies, such as the subsidization of capital could improve efficiency in this case.
Endogenous Growth in an Overlapping Generations Model: The Behavior of Consumers

The change in per capita consumption in the Blanchard-Weil overlapping generations model is given by,

$$\dot{c}(t) = [r(t) - \rho]\dot{c}(t) - n\rho\dot{k}(t)$$

Dividing by per capita income $\dot{y}(t)$, and after we replace the real interest rate from the marginal productivity condition $r(t) = aA - \delta$ and the capital output ratio from the aggregate production function $Y(t) = AK(t)$, we get the following equation for the evolution of the ratio of private consumption to total output,

$$\dot{\bar{c}}(t) = (r(t) - \rho - g)\bar{c}(t) - n\rho\bar{k}(t) = (aA - \delta - \rho - g)\bar{c}(t) - n\rho A^{-1}$$

where, for any variable $X$, we define its ratio to total output $Y$ as $\bar{x} = X/Y$.

From the per capita capital accumulation equation we get,

$$g = (1 - \bar{c}(t))A - n - \delta$$

This suggests that the aggregate endogenous growth rate $g + n$ is equal to the net savings (=investment) rate, times total factor productivity, minus the depreciation rate. Because of constant returns to capital accumulation total factor productivity is equal to the (constant) average and marginal product of capital.

The last two equations can be used to determine the endogenous growth rate and the endogenous savings rate, $1 - \bar{c}(t)$.

$$\bar{c}(t) = \frac{n\rho A^{-1}}{aA - \delta - \rho - g}, \quad g = aA - \delta - \rho - \frac{n\rho}{(1 - n - \delta)A}$$

The endogenous growth rate in an “Arrow-Romer-Blanchard-Weil” model is lower than in the corresponding “Arrow-Romer-Ramsey” model. The reason is that savings, and therefore capital accumulation and growth, are lower in the former model, due to the fact that current generations do not internalize the welfare of future generations.
Endogenous Growth in an Overlapping Generations Model

\[
g = \alpha A - \delta - \rho
g_{R} = \alpha A - \delta - \rho
\]

(c/y) = 0

(c/y)_{E}

E

(c/y)

g

E_{R}

s_{E}

s_{R} = \alpha A - \delta - \rho
Endogenous and Exogenous Growth and Real Convergence

In endogenous growth models of the learning by doing model, there is no convergence process.

In exogenous growth models, any two economies characterized by the same parameters describing the technology of production, household preferences and economic policy, will converge to the same balanced growth path, even if they start from different initial conditions.

In endogenous growth models they will have the same endogenous growth rate, but they will not converge to the same per capita income. Their initial differences will remain for ever.

The available empirical evidence from post war international experience indicates that convergence cannot be dismissed easily.
Inter-temporal Path of Per Capita Income and Real Convergence in Endogenous and Exogenous Growth Models
Models of Human Capital Accumulation and Economic Growth

In the learning-by-doing endogenous growth model, endogenous growth is essentially a by-product of the accumulation of physical capital, because of the assumption that labor efficiency is a function of the aggregate physical capital stock per worker.

An alternative class of growth models (Lucas (1988), Mankiw, Romer and Weil (1992), Jones (2002)) emphasizes the education and training of workers and the accumulation of human capital that it implies. The accumulation of human capital brings about an increase in the efficiency of labor.

Under some conditions, this class of models can also lead to endogenous growth. Endogenous growth in such models is not a by-product of physical capital accumulation, as in the Arrow-Romer model, but also depends on the factors that determine the accumulation of human capital.
The Production Function in Models with Human Capital Accumulation

Firms are competitive, and their production technology is Cobb-Douglas, of the form,

$$Y(t) = AK(t)^\alpha \left( \hat{h}(t)L(t) \right)^{1-\alpha}$$

where, $0 < \alpha < 1$, and $\hat{h}(t)$ denotes the efficiency of labor at instant $t$.

We shall assume that the efficiency of labor is a function of human capital per worker $h$, and exogenous technical progress. This function takes the form,

$$\hat{h}(t) = (h(t))^\gamma \left( e^{gt} \right)^{1-\gamma}$$

where $h$ is human capital per worker, resulting from accumulated education and training, and $0 \leq \gamma \leq 1$; a parameter which measures the relative contribution of human capital in the efficiency of labor. A fraction $1 - \gamma$ of the efficiency of labor depends on technical progress, which grows at an exogenous rate $g$.

Substituting for the efficiency of labor in the production function, we get,

$$Y(t) = AK(t)^\alpha H(t)^\gamma (1-\alpha) \left( e^{gt} L(t) \right)^{(1-\gamma)(1-\alpha)}$$

where $H(t) = h(t)L(t)$ is aggregate human capital.
We assume, generalizing the Solow model, that total household income is either consumed or invested in physical capital or invested in human capital. The percentage of output and income invested in physical capital equals $s_K$ and the percentage invested in human capital equals $s_H$. These percentages are considered exogenous, in the same way that the savings rate is assumed exogenous in the Solow model. The consumption function thus takes the form,

$$C(t) = (1 - s_K - s_H)Y(t)$$

Furthermore we assume that $\delta$, the depreciation rate, is the same for both physical and human capital. With these assumptions, the accumulation equations for physical and human capital are given by,

$$\dot{K}(t) = s_K Y(t) - \delta K(t)$$
$$\dot{H}(t) = s_H Y(t) - \delta H(t)$$

The equilibrium condition in the goods and services market in this case is given by,

$$Y(t) = C(t) + \dot{K}(t) + \dot{H}(t) + \delta \left(K(t) + H(t)\right)$$
The Extended Solow Model in Exogenous Efficiency Units of Labor

In the case where $0 < \gamma < 1$, we can divide the aggregate variables by the number of workers $L$ times the exogenous labor efficiency $e^{g_L}$, and express the model as,

$$
\tilde{y}(t) = A\tilde{k}(t)^\alpha \tilde{h}(t)^{\gamma (1-\alpha)}
$$

$$
\tilde{k}(t) = s_K \tilde{y}(t) - (n + g + \delta)\tilde{k}(t)
$$

$$
\tilde{h}(t) = s_H \tilde{y}(t) - (n + g + \delta)\tilde{h}(t)
$$

$$
\tilde{c}(t) = (1 - s_K - s_H)\tilde{y}(t)
$$

where, $\tilde{y}(t) = \frac{Y(t)}{e^{g_L(t)}}$, $\tilde{c}(t) = \frac{C(t)}{e^{g_L(t)}}$, $\tilde{k}(t) = \frac{K(t)}{e^{g_L(t)}}$, $\tilde{h}(t) = \frac{H(t)}{e^{g_L(t)}}$.

These variables are defined as output, physical capital and human capital, per exogenous efficiency unit of labor.
The Balanced Growth Path in the Extended Solow Model of Exogenous Growth
(Mankiw, Romer and Weil)

On the balanced growth path, all per capita variables grow at the rate of exogenous technical progress $g$. It therefore follows that,

\[
s_K \tilde{y}^* = s_K A \left( \tilde{k}^* \right)^{\alpha} \left( \tilde{h}^* \right)^{\gamma(1-a)} = (n + g + \delta) \tilde{k}^*
\]

\[
s_H \tilde{y}^* = s_H A \left( \tilde{k}^* \right)^{\alpha} \left( \tilde{h}^* \right)^{\gamma(1-a)} = (n + g + \delta) \tilde{h}^*
\]

It thus follows that,

\[
\tilde{k}^* = \frac{s_K}{s_H} \tilde{h}^*
\]

on the balanced growth path, the ratio of physical to human capital is constant, and equal to the ratio of the investment rates in physical and human capital respectively. This follows because equilibrium investment on the balanced growth path is a proportion $n + g + \delta$ of both physical and human capital, per exogenous efficiency unit of labor.

Substituting in the production function and solving, it follows that,

\[
\tilde{h}^* = \left( \frac{A \left( s_K^{\alpha} s_H^{1-a} \right)}{n + g + \delta} \right)^{\frac{1}{(1-\gamma)(1-a)}}, \quad \tilde{k}^* = \left( \frac{A \left( s_K^{1-\gamma(1-a)} s_H^{\gamma(1-a)} \right)}{n + g + \delta} \right)^{\frac{1}{(1-\gamma)(1-a)}}, \quad \tilde{y}^* = \left( \frac{A \left( s_K^{\alpha} s_H^{\gamma(1-a)} \right)}{(n + g + \delta)^{\alpha + \gamma(1-a)}} \right)^{\frac{1}{(1-\gamma)(1-a)}}
\]

Hence, per capita output $\tilde{y}^*(t)$ on the balanced growth path is given by,

\[
\tilde{y}^*(t) = \tilde{y}^* e^{gt} = \left( \frac{A \left( s_K^{\alpha} s_H^{\gamma(1-a)} \right)}{(n + g + \delta)^{\alpha + \gamma(1-a)}} \right)^{\frac{1}{(1-\gamma)(1-a)}} e^{gt}
\]
The Balanced Growth Path in the Extended Solow Model of Exogenous Growth

\[ \hat{y}^*(t) = \hat{y}^* e^{gt} = \left( \frac{A \left( s_K^a s_H^\gamma(1-\alpha) \right)}{(n + g + \delta)^{\alpha+\gamma(1-\alpha)}} \right)^{1/(1-\gamma)(1-\alpha)} e^{gt} \]

The level of per capita output on the balanced growth path depends positively on total factor productivity \( A \) and the shares of output invested in physical and human capital \( (s_K \kappa_{\alpha} s_H) \).

As in the original Solow model it depends negatively on the population growth rate \( n \), the rate of exogenous technical progress \( g \), and the depreciation rate \( \delta \).

The rate of growth of per capita output on the balanced growth path is equal to the rate of exogenous technical progress \( g \).

Thus, in the case where \( 0 < \gamma < 1 \) we have an extended Solow model of exogenous growth. In this model, human capital accumulation plays a role similar to physical capital accumulation in the original Solow model. As the accumulation of human capital causes an increase in the marginal productivity of physical capital and vice versa, the convergence process continues for longer, but eventually the economy converges to a balanced growth path where growth in per capita income is equal to the exogenous rate of technical progress.
Endogenous Growth in the Extended Solow Model
(Mankiw, Romer and Weil)

If $\gamma = 1$, there is no exogenous technical progress. The efficiency of labor only depends on accumulated human capital. The model is this case is an endogenous growth model. Setting $\gamma = 1$ in the aggregate production function we get,

$$\hat{y}(t) = A\hat{k}(t)^{\alpha}\hat{h}(t)^{1-\alpha},$$

where, $\hat{y} = Y/L$, $\hat{k} = K/L$, $\hat{h} = H/L$.

In this model, the balanced growth path would be defined by,

$$\frac{\hat{k}^*(t)}{\hat{h}^*(t)} = \frac{s_K}{s_H}, \quad g = \frac{\hat{y}^*(t)}{\hat{y}(t)} = A\alpha s_K s_H \hat{y}(t)^{1-\alpha} - (n + \delta)$$

On the balanced growth path, the ratio of physical to human capital is stabilized at the ratio of the investment rates in physical and human capital. Because both physical and human capital are growing at the same rate, we have endogenous growth. The accumulation of physical capital causes an increase in income, that in turn causes an increase in human capital through expenditure on education and training. This in turn leads to a further rise in output, which causes new investment in physical capital. The parallel accumulation of physical and human capital leads to endogenous growth.

The endogenous growth rate depends positively on total factor productivity $A$, and a weighted average of the income ratios invested in physical and human capital $s_K$ and $s_H$. The rate of growth of population $n$, and the depreciation rate $\delta$, have a negative impact on the endogenous growth rate.

In all other respects, the properties of the endogenous growth, extended Solow model, are similar to those of the learning by doing model of Arrow and Romer. The main difference is that in this extended Solow model, the accumulation of human capital is not a simple by-product of physical capital accumulation, but the result of explicit investment on human capital, through spending on education and training.
The Jones Model of Human Capital Accumulation

An alternative model, is one in which human capital depends on the time span devoted to education and training. We assume that workers spend part of their time \( \nu \) in education and training, and that this investment of time involves a rate of return \( \psi \). The model is analyzed in Jones(2002).

The production technology is described by the production function,

\[
Y(t) = AK(t)^\alpha \left( \hat{h}(t)L(t) \right)^{1-\alpha}
\]

The efficiency of labor per employee is determined by \( \hat{h}(t) = h(\nu, \psi)e^{\delta t} \).

The function \( h(\nu, \psi) \) determines human capital per worker, as a function of the time span \( \nu \) devoted to education and training, and the rate of return of investment in human capital \( \psi \). A plausible form of this function is the exponential function. The exponential form is plausible because it implies that the proportional change in the function with respect to the time devoted to education and training is equal to the rate of return \( \psi \). We shall therefore assume that,

\[
h(\nu, \psi) = e^{\psi \nu}
\]

In this model, output per worker (per capita output) is given by,

\[
\hat{y}(t) = A\hat{k}(t)^\alpha(e^{\psi \nu}e^{\delta t})^{1-\alpha}
\]

Combining this with the savings assumption of the Solow model, i.e., that there is an exogenous constant savings rate, per capita output on the balanced growth path is determined by,

\[
\hat{y}^*(t) = A \left( \frac{sA}{n + g + \delta} \right)^{1-\alpha} e^{\psi \nu}e^{\delta t}
\]

Per capita output on the balanced growth path is a positive function of the amount of time spent in education and training \( \nu \), and the rate of return on investment in human capital \( \psi \). However, in all other respects, this model is an exogenous growth model, similar to the Solow model.
The Lucas Endogenous Growth Model

An alternative model of investment in human capital is the endogenous growth model of Lucas (1988), which is based, in part, on Uzawa (1965).

Lucas assumed that labor efficiency is equal to human capital per worker, $\hat{h}(t) = h(t)$.

With regard to the production of human capital per worker he assumes that this depends on the existing stock of human capital per worker, times the proportion of non-leisure time that workers devote to education and training. He thus assumes that,

$$\dot{h}(t) = \zeta (1 - u(t))h(t) - (n + \delta)h(t)$$

Only human capital is required for the production of human capital. $u$ is defined as the proportion of the non-leisure time of households that is devoted to the production of goods and services, while $1 - u$ is the proportion of time devoted to education and training. Education and training results in the accumulation of human capital. $\zeta$ is the productivity of human capital in the production of new human capital, $n$ is the population growth rate and $\delta$ the depreciation rate of human capital, assumed for simplicity to be equal to that of physical capital.

On the balanced growth path, $u$ will be constant at $u^*$. Taking the integral of the human capital accumulation equation, we get,

$$h(t) = e^{(\zeta (1 - u^*) - (n + \delta)) t}$$

Lucas assumed that $u$ is chosen optimally by a representative household in order to maximize its intertemporal utility of consumption. As a result, the choice of $u$ depends both on the preferences of the representative household, and on the technological parameters characterizing the production of goods and services and human capital. From the form of the human capital equation, it follows that the endogenous growth rate is equal to $g = \zeta (1 - u^*) - (n + \delta)$, which works like the exogenous rate of technical progress in exogenous growth models. The higher is the steady state proportion of non-leisure time devoted to education and training, the higher the endogenous growth rate. In the Lucas model, $1 - u^*$ is chosen endogenously, and, in conjunction with the exogenous $\zeta$, $\delta$ and $n$, determines the steady state growth rate of per capita output.
Models of the Production of Ideas and Innovations

A final category of growth models, model technical progress as the result of ideas and innovations, that lead to higher total factor productivity or labor efficiency.

These models emphasize the externalities involved in generating new ideas and innovations that increase the efficiency of production. As the learning by doing model of Arrow emphasizes the externalities from capital accumulation, so the ideas and innovations models emphasize the externalities of the production of ideas and innovations.

Although models in this category, sometimes called research and development (R&D) models, date from the late 1960s, the microeconomic foundations of these models and their implications for the functioning of markets were developed in the early 1990s, inspired by the work of Romer (1990). These models, under certain conditions, can lead to endogenous growth as well.
Key Features of Ideas and Innovations

A crucial assumption of such models is that ideas and innovations can improve the efficiency of production, either through total factor productivity or through labor efficiency.

These models recognize that, unlike most other goods and services, the use of an idea and/or innovation by a particular firm, or employee, does not prevent the use of the same idea and innovation from other firms or workers. The use of a particular idea or innovation is non-rivalrous, in contrast to the use of a particular machine or a specific employee. From the time an idea or innovation has been produced, anyone with knowledge of this idea can use it, independently of how many others use it simultaneously. If a firm uses a specific machine or a particular employee, it automatically excludes any other firm from simultaneously using this same machine or this same employee. The property of non-rivalry gives ideas and innovations a character of a quasi public good.

On the other hand, in contrast to purely public goods, the use of an idea can be partially excludable by law. This allows the producer of an idea to charge for the use of his idea. For example, if an idea or innovation is legally covered by a patent, then a firm or an employee who wants to use this idea or innovation will have to pay a fee to the holder of the patent, for their copyright.
Externalities and the Production of Ideas and Innovations

Purely public goods and services are both non rivalrous and non excludable. Other goods and services, such as ideas that can be covered by copyright laws, may be non rivalrous, but may be excludable. Consequently, the producers of goods and services that use an idea or innovation, may be charged for the benefits arising from their use.

The production of non rivalrous and non excludable goods and services implies externalities, which are not reflected in the remuneration of producers of these goods and services. Goods and services that result in positive externalities, such as ideas and innovations, will be produced in smaller quantities than would be socially desirable, and goods and services that result in negative externalities, such as pollution of the environment, will be produced in larger quantities than would be socially desirable.

If the use of ideas and innovations is both non rivalrous and non excludable, then the market will produce fewer ideas and innovations than would be socially desirable. However, if the use of ideas and innovations can be made excludable by the protection of the law on copyright or a patent, then the production of ideas and innovations can rise, as producers of ideas and innovations will be paid for the value of their product.
The Key Elements of an Ideas and Innovations Growth Model

In order to present the key elements of a growth model based on the production of ideas and innovations we start with the production function of goods and services. We assume a Cobb-Douglas production function, of the form,

\[ Y(t) = AK(t)^\alpha \left( H(t)L_y(t) \right)^{1-\alpha} \]

where \( H \) is the existing aggregate stock of ideas and innovations, affecting the efficiency of labor in the goods and services sector, and \( L_y \) is the number of workers who are employed in the goods and services sector.

Apart from goods and services, the economy produces new ideas and innovations. The new ideas and innovations produced per instant are proportional to the number of research workers, i.e., those employed in the production of ideas and innovations. This relationship can be written as,

\[ \dot{H}(t) = h(H(t), L_h(t))L_h(t) \]

where \( L_h \) is the number of workers employed in the production of ideas and innovations, which we shall call research workers or researchers. \( h(\cdot) \) is the average productivity of research workers.

We assume that the average productivity of research workers is a positive function of the total stock of ideas and innovations and a negative function of the number of research workers. The existing stock of ideas and innovations makes every researcher more productive. However, we shall assume that this effect is subject to diminishing returns. On the other hand, as the number of researchers increases, the probability of duplication of effort, that is the probability of simultaneous production of the same idea or innovation by more than one researcher, increases. Thus, the average productivity of research workers falls with the number of researchers. In particular, we will assume that the productivity function of research workers, with the characteristics highlighted above, takes the form,

\[ h(H(t), L_h(t)) = h(t)^\beta L_h(t)^{\gamma-1} \]

where, \( h > 0, 0 < \beta < 1, 0 < \gamma < 1 \).
Diminishing Returns in the Production of New Ideas and Innovations in an Extended Solow Model

With these assumptions, the production function of ideas and innovations can be written as,

$$\dot{H}(t) = hH(t)^\beta L_h(t)^\gamma$$

The production function implies that the existing stock of ideas and innovations has diminishing returns in the production of new ideas and innovations, because the higher the existing stock of ideas and innovations, the more difficult it will be to discover new ones.

It also implies that the number of research workers is also associated with diminishing returns, as the likelihood of duplication of effort in the production of new ideas and innovations is growing with the number of research workers.

The total number of workers in the economy is the sum of workers employed in the production of goods and services and research workers employed in the generation of new ideas and innovations. Consequently,

$$L(t) = L_y(t) + L_h(t)$$

For simplicity we shall assume, as in the Solow model, that,

$$\dot{L}(t) = nL(t), \dot{K}(t) = s_Y Y(t) - \delta K(t), L_h(t) = s_H L(t), L_y(t) = (1 - s_H)L(t)$$
The Determination of the Rate of Technical Progress

Under these assumptions, the production function of new ideas and innovations can be written as,

\[ \frac{\dot{H}(t)}{H(t)} = (h s_H^\gamma) H(t)^{\beta-1} L(t)^\gamma \]

On the balanced growth path, the rate of technical progress will be constant. Taking the logarithm of this equation, and then taking the first derivative of the resulting log linear equation with respect to time, it follows that the rate of technical progress on the balanced growth path will be given by,

\[ g = \frac{\gamma n}{1 - \beta} \]

The rate of technical progress on the balanced growth path, denoted by $g$, will be proportional to the population growth rate $n$. The reason is that the population growth rate determines the growth rate of research workers, who contribute to the generation of new ideas and innovations.

The only other parameters that determine the endogenous rate of technical progress are the parameters of the production function of new ideas and innovations, $\beta$ and $\gamma$. Both have a positive impact on the endogenous rate of technical progress. The higher the elasticity of production of new ideas and innovations with respect to the existing stock of ideas and innovations ($\beta$) and with respect to the number of research workers ($\gamma$), the higher the rate of technical progress on the balanced growth path. Thus, this model attributes technical progress to population growth, and the parameters of the production function of new ideas and innovations.
The Balanced Growth Path in the Ideas and Innovations Solow Model

On the balanced growth path, all per capita variables grow at the endogenous rate of technical progress $g$. Variables per efficiency unit of labor are determined in a way similar to the corresponding Solow model. One can show that, on the balanced growth path, we shall have,

$$
k^* = \left( \frac{s_Y A (1 - s_H)^{1-\alpha}}{n + g + \delta} \right)^{\frac{1}{1-\alpha}}
$$

$$
y^* = A (1 - s_H)^{1-\alpha} (k^*)^\alpha
$$

$$
c^* = (1 - s_Y) y^*
$$

$$
H^*(t) = \left( \frac{(1 - \beta) h}{\gamma n} \right)^{\frac{1}{1-\beta}} \left( s_H L(0) \right)^{\frac{\gamma}{1-\beta}} e^{\left( \frac{\gamma n}{1-\beta} \right)t}
$$

where, $k = K/HL$, $y = Y/HL$, $c = C/HL$, and superscript * denotes the value of the corresponding variable on the balanced growth path.

On the balanced growth path the model behaves exactly as the Solow model, except for the fact that the rate of technical progress $g$ is determined endogenously, and depends on population growth and the parameters that characterize the production function of new ideas and innovations.
Conclusions from the Generalized Growth Models

We have analyzed more general growth models, which, instead of relying only on the assumption of exogenous technical progress, are based on the assumptions of either external effects from the accumulation of physical capital, or investment in human capital, or even the endogenous generation of ideas and innovations.

Endogenous growth models do not necessarily provide for convergence, as the corresponding exogenous growth models. However, the available empirical evidence from post war international experience (see for example Mankiw, Romer and Weil (1992) and Barro (1997)) indicates that the issue of convergence of per capita incomes of the various economies cannot be dismissed easily. This convergence can be explained by generalized models in which there is learning by doing, accumulation of human capital and endogenous technical progress, but not to a degree that would completely neutralize the diminishing returns from the accumulation of physical capital.

Consequently, generalized exogenous growth models, in which there are externalities from capital accumulation, and/or accumulation of human capital and endogenous technical progress, could, in principle, explain most of the aspects of the process of economic growth that cannot be explained by the original Solow model, or the corresponding representative household or overlapping generations models.
Unified Growth Theory and the Transition from Stagnation to Growth

The models we have analyzed in this and the preceding chapters aim to explain the process of economic growth since the take off in economic growth following the industrial revolution. They have little if anything to say about the long centuries of economic stagnation that preceded this period, as we have already documented. They have also very little to say concerning the reasons behind the transition from stagnation to economic growth, and also very little to say about why some countries have been successful in creating the conditions for sustained growth and others have not.

A recent development in the theory of economic growth is the so called unified growth theory. Unified growth theory aims to explain both the process of economic growth after 1820, and the centuries of economic stagnation that preceded it.

Unified growth theory first concentrates on the conditions that prevailed in the so called ‘Malthusian era’, when technological innovations or new resources such as land resulted in an increase in population rather than an increase in per capita income and living standards. Thus, unified growth theory treats population growth as endogenous.

Second, it attempts to explain the transition from stagnation to growth, following the industrial revolution, as a period which, because of the high rate of technological innovation and other social developments, resulted in a demographic transition from high to low population growth rates, allowing for a sustained rise in living standards.

This process was facilitated by social changes which gave emphasis to education and training and the accumulation of human capital.

Unified growth theory combines historical and empirical research with theoretical growth models to explain both the Malthusian process of economic fluctuations that led to fluctuations in population growth rates but more or less stagnant living standards, and the post-Malthusian take off in living standards in the industrial economies.

Unified growth theory does not necessarily clash with or refute the models of economic growth that we have examined. However, it treats them a partial and incomplete accounts of the process of economic growth, and as accounts that are mainly relevant for the post-Malthusian era in the industrial economies. It gives emphasis to the interaction between technical progress and population growth, which are treated as endogenous variables and not exogenous parameters, as in most of the models that we have examined.
Institutions and long-run growth

The models we have presented in this book explain long-run economic growth as a result of the accumulation of physical and human capital and technological change. Why then don't all countries adopt policies that are friendly to such processes. This way, the large differences in per capita income between the developed and less developed world would gradually disappear.

One of the reasons that has been put forward, along with geography and cultural differences, is the role of institutions. The role of institutions has been emphasized by economic historians for a long time, as they have treated the accumulation of physical and human capital and technological progress as only the proximate and not the fundamental determinants of long-run economic growth. As North and Thomas (1973), put it: "the factors we have listed (innovation, economies of scale, education, capital accumulation, etc.) are not causes of growth; they are growth" (p. 2, italics in original). Factor accumulation and innovation are only proximate causes of growth. In North and Thomas's view, and in the view of many other economic historians, the fundamental explanation of comparative growth is differences in institutions. Of primary importance to economic outcomes are economic and political institutions such as civil and property rights, the existence and the imperfections of markets, the nature of the political system etc.

Economic and political institutions are important because they influence the structure of economic incentives. Without the protection of property rights, individuals will not have the incentive to invest in physical or human capital or to adopt more efficient technologies. Economic institutions are also important because they help to allocate resources to their most efficient uses, they determine who gets profits, revenues and residual rights of control. When markets are missing or ignored, potential gains from trade go unexploited and resources are misallocated. Societies with economic institutions that facilitate and encourage factor accumulation, innovation and the efficient allocation of resources will prosper.
Economists also turned their attention to the role of institutions, as exemplified by the pioneering empirical study of Acemoglu et al. (2001) and the literature that is sparked. In recent years, a number of authors have formalized these ideas investigating the effects of institutions on the process of long-run growth. The basic premises of this approach are as follows:

First, economic institutions matter for economic growth because they shape the incentives to invest in physical and human capital and technology, and improve the organization of production. Although cultural and geographical factors may also matter for economic performance, differences in economic institutions are the major source of cross-country differences in economic growth and prosperity. Economic institutions not only determine the aggregate economic growth potential of the economy, but also an array of economic outcomes, including the distribution of resources in the present and the future.

Second, economic institutions are treated as endogenous. They are determined collectively, in large part for their economic consequences. Not all individuals and groups will prefer the same set of economic institutions because different economic institutions lead to different distributions of resources. Consequently, there will typically be conflicts of interest among various groups and individuals over the choice of economic institutions. These conflicts of interest are resolved through the political process.

Third, the distribution of political power in society is also treated as endogenous. Political institutions, similarly to economic institutions, determine the constraints on and the incentives of the key actors in the political sphere. Political power can be either de jure (institutional) or de facto. The de facto power of particular groups depends on their economic resources, which determine both their ability to use (or misuse) existing political institutions and also their option to hire and use force against other groups.
The Persistence of Institutions

There are two sources of persistence in the behavior of such a system:

First, political institutions are durable, and typically, a sufficiently large change in the distribution of political power is necessary to cause a change in political institutions, such as, for example, a transition from dictatorship to democracy.

Second, when a particular group is rich relative to others, this will increase its de facto political power and enable it to push for economic and political institutions favorable to its interests. This will tend to reproduce the initial relative wealth disparity in the future.

Despite these tendencies for persistence, the framework also emphasizes the potential for change. In particular, ‘shocks’, including changes in technologies and the international environment can modify the balance of (de facto) political power in society and can lead to major changes in political institutions and therefore in long-run economic growth.
The New Stylized Facts of Economic Growth

In 1961, Nicholas Kaldor highlighted six ‘stylized’ facts to summarize the patterns that economists had discovered in national income accounts and to check the growth models being developed to explain them. Jones and Romer (2010) attempted at redoing this exercise in order to investigate how much progress has been made. They came up with additional stylized facts that a satisfactory growth model should account for. Their list of additional stylized facts, is the following:

1. Increases in the extent of the market. Increased flows of goods, ideas, finance, and people, via globalization, as well as urbanization, have increased the extent of the market for all workers and consumers.

2. Accelerating growth. For thousands of years, growth in both population and per capita GDP has accelerated, rising from virtually zero to the relatively rapid rates observed in the last century.

3. Variation in growth rates. The variation in the rate of growth of per capita GDP increases with the distance from the technology frontier.

4. Large income and total factor productivity (TFP) differences. Differences in measured inputs explain less than half of the enormous cross-country differences in per capita GDP.

5. Increases in human capital per worker. Human capital per worker is rising dramatically throughout the world.

6. Long-run stability of relative wages. The rising quantity of human capital, relative to unskilled labor, has not been matched by a sustained decline in its relative price.
Whereas Kaldor's original facts were accounted for almost entirely using the neoclassical growth model, the facts highlighted by Jones and Romer reveal the broader reach of modern growth theory. To capture these facts, a growth model must consider the interaction between ideas, institutions, population, and human capital.

Two of the major facts of growth, its acceleration over the very long run and the extraordinary rise in the extent of the market associated with globalization, are readily understood as reflecting the defining characteristic of ideas, their non-rivalry.

The next two major facts, the enormous income and total factor productivity differences across countries, as well as the stunning variation in growth rates for countries far behind the technology frontier, testify to the importance of institutions and institutional change.

The final two facts of Jones and Romer parallel two of Kaldor's original observations, but while his emphasis was on physical capital, the emphasis in modern growth theory is on human capital. Human capital per worker is rising rapidly, and this occurs despite no systematic trend in the wage premium associated with education.

According to Jones and Romer, these facts also reveal important complementarities among the key endogenous variables.

The virtuous circle between population and ideas accounts for the acceleration of growth.

Institutions may have their most important effects on cross-country income differences by hindering the adoption and utilization of ideas from throughout the world. Institutions like public education and the university system are surely important for understanding the growth in human capital. And institutions are themselves based on ideas - inventions that shape the allocation of resources - and the search for better institutions is unending.

Finally, the rising extent of the market, which raises the return to ideas and to the human capital that is a fundamental input to the production of ideas, may help explain why the college wage premium has not fallen systematically, despite the huge increases in the ratio of college graduates to high school graduates.
Notes on the Literature

For comprehensive surveys of the extensive recent literature on economic growth and a number of additional models, including political economy models, see *The Handbook of Economic Growth*, edited by Aghion and Durlauf (2005) and Aghion and Durlauf (2014). There are also excellent presentations in books by Acemoglu (2009), Aghion and Howitt (2009), and Barro and Sala-i-Martin (2004). On unified growth theory see Galor (2011).