Endogenous Growth and Cross-Country Income Differences

By Peter Howitt*

A multi-country Schumpeterian growth model is constructed. Because of technology transfer, all R&D-performing countries converge to parallel growth paths. All other countries stagnate. Any parameter change that would have raised a country's growth rate in standard Schumpeterian theory will permanently raise its productivity and per-capita income relative to other countries and raise the world growth rate. Transitional dynamics are analyzed for each country and for the world economy. Steady-state income differences obey the same equation as in neoclassical theory, but since R&D is positively correlated with investment rates, capital accumulation accounts for less than estimated by neoclassical theory. (*JEL* E10, O40)

Cross-country evidence on income differences has been used in recent years to cast doubt on endogenous growth theory. N. Gregory Mankiw, David Romer and David N. Weil (1992) argue that the neoclassical Solow-Swan growth model with exogenous technological progress and diminishing returns to capital explains most of the cross-country variation in output per person. Paul Evans (1996) shows that the dispersion of per-capita income across advanced countries has exhibited no tendency to rise over the postwar era, as would be predicted by some endogenous growth models; instead, these countries have been converging[1] to parallel growth paths of the sort implied by the Solow-Swan model with a common world technology. Similarly, it is often argued that the evidence of conditional β-convergence coming from cross-country growth regressions (see Robert J. Barro and Xavier Sala-i-Martin, 1995) is consistent with neoclassical theory but not with endogenous growth theory.

The present paper challenges the interpretation that these arguments place on the facts, by arguing that the evidence is actually more supportive of the Schumpeterian version of endogenous growth theory than of neoclassical theory. The paper constructs a multi-country endogenous growth model that combines elements of the Solow-Swan model and the creative destruction model of Aghion and Howitt (1992). The model is consistent not only with the critics' evidence but also with other evidence that the Solow-Swan model cannot account for.[2]

The reason why the Schumpeterian model can account for more than the neoclassical model is that per-capita income varies across countries not only because of differences in capital stocks per worker but also because of differences in productivity. Empirical research by Daniel Trefler (1993, 1995), Robert E. Hall and Jones (1999) and others shows that these productivity differences are substantial, even among advanced countries. The Schumpeterian approach offers an explanation for productivity differences, while the neoclassical approach offers none.

To account for convergence, the model follows the lead of others[3] who have argued that convergence takes place not only through diminishing returns to capital but also through technology transfer. In the model presented below, countries are connected by R&D spillovers of the sort that David T. Coe and Elhanan Helpman (1995) and Eaton and Kortum (1996) estimate to be substantial.[4] The model implies that all countries engaging in R&D will grow at the same rate in the long run. Convergence is restricted however to this select group of countries. Those in which there is not a strong enough incentive to perform R&D will not grow at all in the long run.

Unlike the neoclassical model, the Schumpeterian model implies that even those countries that converge to a common positive growth rate will have different productivity levels. Long-run differences in productivity are endogenous and depend on the incentives to innovate and to accumulate capital. Any parameter change that would raise a country's long-run growth rate if it operated in isolation from the rest of the world will raise its long-run relative productivity level in the multi-country model, and will have a (possibly small) positive effect on the long-run world growth rate.[5]

Because the model is characterized by a standard Cobb-Douglas aggregate production function in each country, it implies the same "steady-state" equation that Mankiw, Romer and Weil use to explain cross-country differences in per-capita income. However, it also implies that their identifying assumption that productivity differences across countries are uncorrelated with investment rates is generally invalid, and that this leads them to overestimate the impact of increased capital on a country's steady-state level of per-capita income. Specifically, if R&D intensities are positively correlated with steady-state capital stocks, as is the case in the Mankiw-Romer-Weil sample, then the Cobb-Douglas coefficient on capital, which would be capital's share of GDP in a general competitive equilibrium, is strictly less than the value estimated by their regression.

This capital coefficient is a key variable in the debate over endogenous growth. Mankiw, Romer and Weil argue that their finding of a coefficient less than unity refutes the AK version of endogenous growth. But they have to augment the neoclassical model, by interpreting capital to include human as well as physical capital, in order to reconcile their coefficient estimate with direct measures of capital's share of GDP. The present paper shows that the Schumpeterian version of endogenous growth theory, instead of implying a coefficient of unity, is actually less in need of augmentation than neoclassical theory, since it implies that the coefficient is less than the value estimated under the maintained hypotheses of neoclassical theory.

Proponents of neoclassical theory argue further that the coefficient value estimated by Mankiw, Romer and Weil (about 0.7) is consistent with estimated rates of convergence; the Solow-Swan model with this coefficient value would imply a convergence rate approximately equal to the frequently estimated 2 percent per year. However, the Schumpeterian model below implies that neoclassical theory misspecifies the relationship between capital's coefficient and the rate of convergence. It implies that the two parameters will be negatively related to each other, as in neoclassical theory, but that, given any coefficient value, convergence will be slower than predicted by the neoclassical model.

This is because as an economy approaches its steady state from below, its productivity will be rising relative to the rest of the world, which will offset the growth-dampening effects of a diminishing marginal product of capital. Neoclassical theory, by neglecting endogenous movements in productivity, thus omits a factor that serves to attenuate the convergence process. A calibration exercise below shows that a 2 percent convergence rate is consistent with a capital coefficient of no more than 0.6 when the corresponding neoclassical model would imply a coefficient of 0.7. Thus the smaller coefficient that Schumpeterian theory infers from the Mankiw-Romer-Weil steady-state regression is also consistent with existing estimates of convergence rates.

The basic production relations underlying the model are spelled out in Section I. Section II presents a brief account of endogenous innovations by profit-maximizing R&D firms, along the lines of Aghion and Howitt (1992). Section III describes the process of technology transfer, showing how it leads to convergence, and derives a pair of differential equations characterizing an economy taking as given the world rate of technological progress. One of these equations is identical to the fundamental differential equation of the Solow-Swan model, except that the rate of labor-augmenting technological progress here is endogenous; the other equation describes the evolution of the economy's productivity level relative to the rest of the world. Section IV analyzes the economy's steady state and characterizes the determination of the country's relative productivity and relative income per person. Section V analyses transitional dynamics and discusses "club convergence." Section VI considers the dynamic evolution of a global system of several different economies, each describable by the pair of differential equations derived in Section III, and shows how this system determines the world rate of technological progress. Section VII uses the model to interpret empirical findings with respect to cross-country income differences. Section VIII concludes.

I. Production Relations

Consider a single country in a world economy with m different countries. There is one final good, produced under perfect competition by labor and a continuum of intermediate products, according to the production function:

$$Y_t = \int_0^{N_t} A_t(i) F(x_t(i), L_t/N_t)\, di, \tag{1}$$

where Y_t is the country's gross output at date t, L_t is the flow of labor used in production, N_t measures the number of different intermediate products produced and used in the country, $x_t(i)$ is the flow output of intermediate product $i \in [0, N_t]$, $A_t(i)$ is a productivity parameter attached to the latest version of intermediate product i, and $F(\cdot)$ is a smooth, concave, constant-returns production function. For simplicity attention is restricted to the Cobb-Douglas case:

$$F(x, \ell) \equiv x^\alpha \ell^{1-\alpha},\ 0 < \alpha < 1. \tag{2}$$

To focus on technology transfer as the main connection between countries, assume that there is no international trade in goods or factors. Each intermediate product is specific to the country in which it is used and produced although, as we shall see, the idea for how to produce it generally originates in other countries.

Labor supply and population size are identical. They grow exogenously at the fixed proportional rate g_L. The number of products grows as a result of serendipitous imitation, not deliberate innovation.[6] Imitation is limited to domestic intermediate products; thus each new product will have the same productivity parameter as a randomly chosen existing product within the country. Each person has the same propensity to imitate $\xi > 0$. Thus the aggregate flow of new products is:

$$\dot{N}_t = \xi L_t.$$

Since the population growth rate g_L is constant, the number of workers per product L_t/N_t converges monotonically to the constant:[7]

$$\ell = g_L/\xi. \tag{L}$$

Assume that this convergence has already occurred, so that $L_t = \ell N_t$ for all t.

The form of the production function (1) ensures that growth in product variety does not affect aggregate productivity. This and the fact that population growth induces product proliferation guarantees that the model does not exhibit the sort of scale effect that Jones (1995) argues is contradicted by postwar trends in R&D spending and productivity.[8] That is, a bigger population will not by itself raise the incentive to innovate by raising the size of market that can be captured by an innovator, because each innovation is restricted to a single intermediate product, and the number of buyers per intermediate product does not increase with the size of population.[9]

Final output can be used interchangeably as a consumption or capital good, or as an input to R&D. Each intermediate product is produced using capital, according to the production function:

$$x_t(i) = K_t(i)/A_t(i), \tag{3}$$

where $K_t(i)$ is the input of capital in sector i. Division by $A_t(i)$ in (3) indicates that successive vintages of the intermediate product are produced by increasingly capital-intensive techniques.[10]

Innovations are targeted at specific intermediate products. Each innovation creates an improved version of the existing product, which allows the innovator to replace the incumbent monopolist until the next innovation in that sector.[11] The incumbent monopolist of each product operates with a price schedule given by the marginal product: $p_t(i) = A_t(i)\,\alpha\,(x_t(i)/\ell)^{\alpha-1}$ and a cost function equal to $(r_t + \delta)\,A_t(i)\,x_t(i)$, where r_t is the rate of interest and δ is the fixed rate of depreciation.

Since each intermediate firm's marginal revenue and marginal cost schedules are proportional to $A_t(i)$, and since firms differ only in their value of $A_t(i)$, they all choose to supply the same quantity of intermediate product: $x_t = x_t(i)$ for all i. Putting this common quantity into (3), and assuming that the total demand for capital equals the given supply K_t, yields:

$$x_t(i) = x_t = k_t\ell, \tag{4}$$

where k_t is the capital stock per "effective worker" K_t/A_tL_t, and A_t is the average productivity parameter across all sectors.[12]

Substituting from (4) into (1) and (2) shows that output per effective worker is given by a familiar Cobb-Douglas function of capital per effective worker:

$$Y_t/L_tA_t = k_t^{\alpha} \equiv f(k_t).$$

Substituting from (4) into the standard profit-maximization condition of each intermediate firm, and using the above definition of $f(\cdot)$, yields the equilibrium interest rate:

$$r_t = \alpha f'(k_t) - \delta, \tag{5}$$

and shows that each local monopolist will earn a flow of profits proportional to its productivity parameter $A_t(i)$, namely:

$$\pi_t(i) = A_t(i)\,\alpha\,(1-\alpha)\,k_t^{\alpha}\ell \equiv A_t(i)\,\widetilde{\pi}_t(k_t)\,\ell. \tag{6}$$

II. Innovations

Innovations result from domestic R&D that uses technological knowledge coming from all over the world. That is, at any date there is a world-wide "leading-edge technology parameter:"

$$A_t^{\max} \equiv \max\left\{A_{jt}(i) \ \mid i \in [0, N_{jt}]\,,\ j = 1, ..., m\right\},$$

where the j subscript denotes a variable specific to country j. Each innovation in sector i of a country at date t results in a new generation of that country's i^{th} product, whose productivity parameter equals[13] $A_t^{\max}$.

The Poisson arrival rate ϕ_t of innovations in each sector is:

$$\phi_t = \lambda n_t;\ \ \lambda > 0,$$

where λ is a parameter indicating the productivity of R&D, and where n_t is the productivity-adjusted quantity of final output devoted to R&D in each sector; i.e. R&D expenditure per intermediate product, divided by $A_t^{\max}$. The division by $A_t^{\max}$ takes into account the force of increasing complexity; as technology advances, the resource cost of further advances increases proportionally.[14]

Suppose that R&D expenditures are subsidized at the proportional rate $\psi < 1$. The subsidy rate ψ is a proxy for all distortions and policies that impinge directly on the incentive to innovate. It can be negative, in which case the distortions and policies favoring innovation are outweighed by those discouraging it.

The arbitrage condition governing the level of R&D is that the net marginal cost of R&D $(1-\psi)$ equal the marginal effect $\lambda/A_t^{\max}$ of R&D on the arrival rate times the expected discounted value of the flow of profits that a successful innovator will earn. Hence the analogue to the research-arbitrage equation of Aghion and Howitt (1992):[15]

$$1 - \psi = \lambda \frac{\widetilde{\pi}(k_t)\,\ell}{r_t + \lambda n_t}. \tag{7}$$

The discount rate applied in (7) is the rate of interest plus the rate of creative destruction λn_t; the latter is the instantaneous flow probability of being displaced by an innovation.

Equations (L), (5) and (7) can be solved for the country's R&D intensity n_t at any date as a function of the capital intensity k_t and the country-specific parameters:[16]

$$\theta \equiv (\lambda, g_L, \psi) .$$

Assume provisionally that the resulting R&D intensity is positive. It can be written as:

$$n_t = \widetilde{n}\,(k_t; \theta)\,,\ \ \widetilde{n}_k > 0,\, \widetilde{n}_\lambda > 0,\, \widetilde{n}_{g_L} > 0,\, \widetilde{n}_\psi > 0.$$

This relationship is invariant to the global productivity parameter $A_t^{\max}$, because both the cost and the reward to R&D are proportional to $A_t^{\max}$. As explained in Howitt and Aghion (1998), an increase in the capital intensity k induces more R&D by raising the reward to innovation, which is proportional to aggregate output, and by reducing the interest rate used for discounting that reward. A faster rate of population growth g_L induces more R&D through a "scale effect," by increasing the equilibrium number ℓ of people per product.

III. Productivity Growth and Capital Accumulation

A country's average productivity parameter A_t grows as a result of innovations, each of which replaces the pre-existing productivity parameter $A_t\,(i)$ in a sector by the worldwide leading-edge parameter $A_t^{\max}$. The rate of increase in this average equals the flow rate of innovation λn_t times the average increase in $A_t\,(i)$ resulting from each innovation.[17] Since innovations are uniformly distributed across all sectors, this means:

$$\dot{A}_t = \lambda n_t\,(A_t^{\max} - A_t)\,. \tag{8}$$

If the leading-edge parameter $A_t^{\max}$ were to remain unchanged then according to (8) each country's average productivity level would converge to $A^{\max}$, as long as λn was positive. But if the leading edge is constantly increasing then a country with a higher level of innovation will eventually have an average productivity level that is permanently closer to $A_t^{\max}$, because a larger fraction of its sectors will have experienced a recent innovation embodying the leading-edge technology. In short, more innovative economies will be more productive because their intermediate products are generally more up-to-date.

More precisely, suppose that the world rate of technological progress at any given date t is:

$$g_t \equiv \dot{A}_t^{\max}/A_t^{\max}.$$

If the ratio $\lambda n_t/g_t$ is a positive constant, then the reasoning of Aghion and Howitt (1998, pp.115-16) shows that the cross-sectional distribution of productivity parameters within the country will converge asymptotically to a power-law distribution:

$$\Phi_t\left(A\right) \equiv \text{Fraction of sectors with } \left\{A_t\left(i\right) \leq A\right\} = \left(A/A_t^{\max}\right)^{\lambda n/g}.$$

Thus, as shown in Figure 1, a country with $\lambda n = g$ will come to have a uniform distribution of productivity parameters between 0 and $A_t^{\max}$, while one with a higher rate of innovation will have a convex distribution (with more sectors close to $A_t^{\max}$) and one with a lower rate of innovation will have a concave distribution (with more sectors close to 0). More generally, in the long run a country's distribution will stochastically dominate that of any other country with a lower innovation rate.

Figure 1 here

Let $a_t \equiv A_t/A_t^{\max}$ denote the country's average productivity relative to the leading edge. It follows from (8) and the definition of g_t that:

(9) $$\dot{a}_t = \lambda\widetilde{n}\left(k_t;\theta\right)\left(1-a_t\right) - a_t g_t.$$

Equation (9) contains the mechanism through which technology transfer makes the country's productivity-growth rate converge to the global growth rate. An increase in R&D will cause productivity growth to rise temporarily, but as the gap $(1-a_t)$ narrows between its average productivity and the global leading-edge technology, innovations will raise average productivity by less and less, and this will slow down the growth rate of the average. Thus given any fixed R&D intensity n and any global growth rate g, the country's relative productivity a will converge to the steady-state level $\lambda n / (g + \lambda n)$. Accordingly, all countries' growth rates will converge to the same value but their average productivity levels will not; countries with a larger rate of innovation will also have permanently higher average productivity.

Assume that the investment rate $\left(\dot{K} + \delta K\right)/Y$ is a constant s. Since $k = K/AL$, it follows from (8) that:

(10) $$\dot{k}_t = sk_t^{\alpha} - \left[\delta + g_L + \lambda \widetilde{n}\left(k_t; \theta\right)\left(a_t^{-1} - 1\right)\right] k_t.$$

Equation (10) is the usual differential equation of neoclassical growth theory, except that rate of technological progress on the right hand side is now endogenous. Since this rate converges to the world rate g in the long run, the steady-state capital intensity will therefore be identical to that of neoclassical growth theory.

The two differential equations (9) and (10) constitute a two-dimensional dynamical system governing the behavior of a country's relative productivity a_t and capital stock per effective worker k_t. Together with initial values a_0 and k_0 and the trajectory of world productivity growth $\{g_t\}_0^{\infty}$ they characterize completely the evolution of the economy.

IV. Steady-State Analysis

Suppose that the world growth rate g is constant. Consider an economy in a steady state, i.e. with constant relative productivity a and constant capital per effective worker k. According to (9) and (10) the steady state is defined by the two equations:

(A) $$a = \frac{\lambda \widetilde{n}\left(k; \theta\right)}{g + \lambda \widetilde{n}\left(k; \theta\right)},$$

(K) $$sk^{\alpha - 1} = \delta + g_L + g.$$

These equations can be solved recursively. Equation (K) is the same equation that determines k in neoclassical theory. Its solution can be substituted into (A) to solve for a.

Assume that the steady-state rate of interest $r = \alpha^2 k^{\alpha-1} - \delta$ is positive. Straightforward comparative-statics techniques yield:

Proposition 1 *A country's steady-state relative productivity a depends positively on its investment rate s, the productivity of its R&D λ and its R&D subsidy rate ψ, and negatively on the world growth rate g.*

An increase in λ would raise a by inducing a rise in R&D intensity, and also by making R&D more productive; on both counts there would be a temporary rise in productivity growth above the world rate g, which would eventually be choked off as the country came closer to the global technology frontier. The result would be a permanent rise in the country's relative productivity. An increase in ψ would affect relative productivity analogously. An increase in s or a decrease in g would raise the capital intensity k, exactly as in neoclassical growth theory. This rise in k would induce more R&D, which again would result in an increase in the country's relative productivity. An increase in population growth would have an ambiguous effect on relative productivity because although it would have a positive direct effect on the R&D function $\widetilde{n}$, it would also result in a lower capital intensity through the normal neoclassical mechanism embodied in equation (K).

According to Proposition 1, the determinants of an open economy's relative productivity are precisely the same as the determinants of the growth rate of the analogous closed economy. For suppose we closed the model, as in the analysis of Howitt and Aghion (1998), by assuming that the technology parameter $A_t^{\max}$ embodied in each new innovation was the leading edge within that country instead of within the whole world, and that the growth rate of $A_t^{\max}$ was proportional to the flow of innovations λn_t within the country. Then the growth rate in a steady state would be the solution to:

$$g = \sigma \lambda \widetilde{n} \left(\widetilde{k}(s, \theta, g), \theta \right), \tag{CG}$$

where $\sigma > 0$ is a spillover coefficient and $\widetilde{k}(s, \theta, g)$ is the solution to the neoclassical steady-state equation (K). Comparison of (CG) with (A) shows that any parameter change that would raise a in the open economy would raise g in the closed economy, and vice-versa.

Let $y_t = Y_t / L_t A_t^{\max} = a_t k_t^{\alpha}$ denote the country's per-capita income, relative to global productivity. Since anything that raises either the country's relative productivity a_t or its capital stock per effective worker k_t will raise its relative per-capita income, we have:[18]

Proposition 2 *In a steady state, a country's relative per-capita income y depends positively on its investment rate s, the productivity of its R&D λ and its R&D subsidy rate ψ, and negatively on the world growth rate g.*

The model thus provides a more complete account of cross-country differences in per-capita income than does neoclassical growth theory. For these differences depend on differences in capital per effective worker and differences in productivity. The present Schumpeterian approach offers an explanation for both sets of differences whereas the neoclassical approach must take the latter as given.

Although our assumption of an exogenous investment rate s affects the dynamics of convergence, it does not affect the analysis of steady states. For example, if we were to allow investment to be governed by Ramsey-style intertemporal utility maximization with a constant rate of time preference and an iso-elastic instantaneous utility function, the investment rate would still be constant in a steady state. Equations (A) and (K) would still apply, and Propositions 1 and 2 would still hold, with the effects of s being interpreted as the effects of corresponding changes in the parameters of the intertemporal utility function.

Likewise, the assumption of no trade in final goods is not crucial for the steady-state analysis. If we were to allow international trade in (identical) final goods, and assume perfect capital mobility, each country would now take as given not only the growth rate g but also the rate of interest r. To avoid all countries having the same investment rate we could take into account the factor stressed by Chaari, Kehoe and McGrattan (1997); namely distortions to investment. Thus the capital intensity in each country would be $k(r, \psi_k)$, the solution to the modified equilibrium condition:

$$r = \alpha f'(k) - \delta + \psi_k,$$

where ψ_k is the subsidy rate to capital. A slight modification of the analysis of Barro, Mankiw and Sala-i-Martin (1995) shows that equations (A) and (K) would still apply, with s now endogenous. Propositions 1 and 2 would still hold, with the effect of an increase in s being replaced by the effect of a decrease in the world interest rate r or an increase in the capital subsidy rate ψ_k.

V. Transitional Dynamics and "Club Convergence"

Figure 2 below represents the phase diagram of equations (9) and (10). It illustrates the non-steady-state dynamics of an economy that takes as given a constant world growth rate g. From any initial position the economy converges to its steady state along a non-cyclic path that converges to the upward-sloping turnpike DD.

Figure 2 here

Suppose a steady state were disturbed by an increase in the investment rate s. As shown in Figure 3(a), the $\dot{k} = 0$ locus would shift up, and both variables would rise monotonically to the new steady state. The ultimate increase in capital per effective worker would be exactly the same as in neoclassical growth theory, but the fact that relative productivity also rises means that the ultimate increase in relative per-capita income would be strictly greater than in neoclassical theory.

Figure 3 here

An increase in the productivity of R&D λ or the R&D subsidy rate ψ would shift both of the stationary loci to the right by the same amount, as shown in Figure 3(b). Relative productivity would rise monotonically to its new equilibrium value, while the capital stock per effective worker would first fall and then rise back to its original value. In the end, the country's relative per-capita income would have risen by the same proportional amount as its relative productivity.

The present approach allows for the possibility of "club convergence." We have until now assumed that the equilibrium level of R&D is positive; i.e. that the solution $\widetilde{n}$ to the research-arbitrage equation (7), and hence the steady-state relative productivity a determined by (A), are strictly positive. However, for low enough values of the subsidy rate, the investment rate, or the productivity of R&D, there will be no such solution when the capital intensity has adjusted to its steady-state value. To accommodate this, we now replace (7) by the more general Kuhn-Tucker condition:

$$1 - \psi \geq \lambda \frac{\widetilde{\pi}(k_t)\,\ell}{r_t + \lambda n_t},\ n_t \geq 0, \text{ with at least one inequality.}$$

This generalization makes no difference to the analysis except for the interpretation of the steady state, which is still governed by equations (A) and (K), where $\widetilde{n}(\cdot)$ is the solution to the generalized research-arbitrage equation. Countries for which $n > 0$ in the steady state have $a > 0$, and grow at the world rate g. Countries for which $n = 0$ in the steady state have $a = 0$ and do not grow at all.[19] Thus the analysis contains not only an account of why convergence takes place between members of the club but also why some countries have not joined the club.[20]

VI. Growth of the World Economy

The growth rate g_t of the world's leading-edge technology parameter $A_t^{\max}$ is determined by a spillover process that constitutes part of the mechanism of technology transfer (the other part being the use of $A_t^{\max}$ by innovators in every country). That is, the global technology frontier expands as a result of innovations everywhere, which produce knowledge that feeds into R&D in other sectors and in other countries.

Since population grows in all countries, so does the number of intermediate products N_t. Thus the aggregate flow of innovations in a country, $N_t \phi_t$ grows steadily even in a steady state. Suppose that as the number of products grows, the marginal contribution of each innovation to global knowledge falls proportionally, reflecting the increasingly specialized nature of the knowledge resulting from the innovation. That is, suppose that:[21]

$$\text{(S)} \qquad g_t \equiv \dot{A}_t^{\max}/A_t^{\max} = \sum_{j=1}^{m} \left(\sigma_j/N_{jt}\right) N_{jt}\phi_{jt} \equiv \sum_{j=1}^{m} \sigma_j \lambda_j n_{jt},$$

where the spillover coefficients σ_j are all non-negative.[22]

The global spillover equation (S) is what links the different countries. Substituting it into equations (9) and (10) for each country results in the $2m$-dimensional system:

$$\text{(11)} \qquad \dot{a}_{jt} = \lambda_j \widetilde{n}\left(k_{jt};\theta_j\right)\left(1-a_{jt}\right) - a_{jt}\sum_{q=1}^{m}\sigma_q\lambda_q\widetilde{n}\left(k_{qt};\theta_q\right);\ j=1,...,m,$$

$$\text{(12)} \qquad \dot{k}_{jt} = s_j k_{jt}^{\alpha} - \left[\delta + g_{Lj} + \lambda_j\widetilde{n}\left(k_{jt};\theta_j\right)\left(a_{jt}^{-1}-1\right)\right]k_{jt};\ j=1,...,m.$$

A steady state for the world economy is a rest point of this system. The steady-state world growth rate is the solution to:[23]

$$\text{(WG)} \qquad g = \sum_{j=1}^{m}\sigma_j\lambda_j n\left(\widetilde{k}\left(s_j,\theta_j,g\right);\theta_j\right).$$

Equation (WG) is a straightforward generalization of the steady-state growth equation (CG) for the closed economy. Comparison of these two equations shows that any change in a country-specific parameter that would have raised the growth rate in that economy if it were closed will have a (possibly small) positive effect on the world growth rate when the economy is part of a global system with technology transfer. Specifically:

Proposition 3 *The steady-state world growth rate g depends positively on each country's investment rate s_j, productivity of R&D λ_j and R&D subsidy rate ψ_j.*

The global stability of the international system (11)$\sim$(12) is an open question. However, Appendix B shows that the steady state exhibits local asymptotic stability.

A special case is that in which only one country, the world technology leader, generates positive spillovers ($\sigma_1 > 0$, $\sigma_j = 0$ for $j = 2, ..., m$). In this case, the system would be dichotomized. The leader would behave as a closed economy, unaffected by R&D in the rest of the world, with a long-run growth rate determined by the steady-state equation (CG) above, while all other countries would behave according to (9) and (10) with g equal to the leader's productivity-growth rate.

Except in this limiting case, however, even the leading country will be affected by technological discoveries in the rest of the world. If $\sigma_j > 0$ for some lagging countries, then an acceleration in their pace of innovation would cause the leading-edge technology parameter $A_t^{\max}$ to accelerate, and growth would eventually rise even in the leading country. Under this scenario the rise in growth in the leading country would lag the rise in the lagging countries.

That is, the growth rate of a country's per-capita income Y_{jt}/L_{jt} at any date is:

$$G_{jt} = \alpha \dot{k}_{jt}/k_{jt} + \dot{A}_{jt}/A_{jt}. = \alpha \dot{k}_{jt}/k_{jt} + \dot{a}_{jt}/a_{jt} + g_t.$$

From this and the dynamic equations (11) and (12):

$$\text{(G)} \qquad G_{jt} = \alpha \left(s_j k_{jt}^{\alpha-1} - \delta - g_{Lj} \right) + (1-\alpha) \lambda_j \widetilde{n} \left(k_{jt}; \theta_j \right) \left(a_{jt}^{-1} - 1 \right).$$

Thus if a group of lagging countries were to experience a change in the parameter vectors θ_j that raised their R&D intensities $\widetilde{n}_j$, their growth rates would immediately rise. But since those parameters do not enter directly into the leader's growth equation (equation (G) for $j = 1$), the leader would not experience a rise in growth until the acceleration in global technological growth had reduced its relative productivity level a_{1t}, at which point it too would begin to grow faster as a result of technology transfer.

VII. Interpreting Evidence on Income Differences

A. Cross-Country Regressions

Rewriting equation (K) in logs and using the fact that $Y_t/L_t A_t^{\max} = a_t k_t^{\alpha}$, we arrive at the familiar steady-state equation:

$$\ln\left(\frac{Y_t}{L_t}\right) = \frac{\alpha}{1-\alpha}\left[\ln s - \ln\left(\delta + g_L + g\right)\right] + \ln A_t^{\max} + \ln \widetilde{a}\left(s, \theta, g\right)$$

where $\widetilde{a}$ is the steady-state value of the country's relative productivity as characterized by Proposition 1 above. Suppose we allow for physical and human capital to be accumulated according to separate investment rates s_k and s_h, and generalize the production function (3) to make it Cobb-Douglas with constant returns in human and physical capital. Then going through exactly the same argument as above we arrive at the analogous steady-state equation:

$$\begin{aligned}(13) \quad \ln\left(\frac{Y_t}{L_t}\right) &= \frac{\alpha}{1-\alpha-\beta}\left[\ln s_k - \ln\left(\delta + g_L + g\right)\right] \\ &\quad + \frac{\beta}{1-\alpha-\beta}\left[\ln s_h - \ln\left(\delta + g_L + g\right)\right] + \ln A_t^{\max} + \ln \widetilde{a}\left(s_k, s_h, \theta, g\right)\end{aligned}$$

where α and β are respectively the coefficients of the two kinds of capital. This is the equation estimated by Mankiw, Romer and Weil (1992), who do not however include the relative productivities $\widetilde{a}$, but assume instead that these are uncorrelated with the other regressors in the equation.

Under this identifying assumption Mankiw, Romer and Weil show that the equation accounts for over three quarters of the variation in per-capita income in their sample of

98 countries, with estimated coefficients α and β that sum to about 0.7. Since the only explanatory variables in this equation are proxies for steady-state per-capita capital stocks, they conclude that capital accumulation alone is sufficient to account for most of the variation in per-capita income across countries, that the neoclassical growth model with exogenous technological progress is thereby vindicated, and that the AK version of endogenous growth, according to which $\alpha + \beta = 1$, is thereby rejected.

The fact that the same equation can be derived from the present Schumpeterian endogenous growth framework implies that its empirical success does not refute endogenous growth theory in general. Moreover, the present analysis suggests that the identifying assumption to the effect that productivity is uncorrelated with steady-state per-capita capital stocks is wrong, and that therefore the results of Mankiw, Romer and Weil need to be re-examined.[24] Specifically, countries with higher steady-state per-capita capital stocks will have higher R&D intensities, through the mechanism discussed above in connection with Proposition 1. This yields a presumption that R&D intensities will be positively correlated with steady-state capital intensities. If so, then according to (A) the error term in the equation that omits productivity differences will be positively correlated with the included explanatory variables, leading to upward biased coefficients on net investment and hence an upward biased measure of the contribution of capital to economic growth.[25] Some of the explanatory power that this equation attributes to differences in equilibrium capital/labor ratios should instead have been attributed to differences in productivity levels.

Some idea of the size of the resulting bias could be inferred by adding the investment rate in R&D to the Mankiw-Romer-Weil regression. Taking a log-linear approximation of (A), and recalling that by definition n is the level of R&D per intermediate product, normalized by the leading-edge technology $A_t^{\max}$, we have:

$$\begin{aligned}\ln a &\simeq \text{constant } + \kappa \ln n \\ &= \text{constant } + \kappa \left(\ln \left(\frac{R\&D}{Y} \right) + \ln \left(\frac{Y}{L} \right) + \ln \ell - \ln A_t^{\max} \right).\end{aligned}$$

Substituting (L) into this equation, and substituting the result into the steady-state equation (13) yields:

$$\begin{aligned}\ln\left(\frac{Y_t}{L_t}\right) &= \frac{1}{1-\kappa}\frac{\alpha}{1-\alpha-\beta}\left[\ln s_k - \ln\left(\delta+g_L+g\right)\right] + \\ &\quad \frac{1}{1-\kappa}\frac{\beta}{1-\alpha-\beta}\left[\ln s_h - \ln\left(\delta+g_L+g\right)\right] + constant + \\ &\quad \frac{\kappa}{1-\kappa}\ln\left(\frac{R\&D}{Y}\right) + \frac{\kappa}{1-\kappa}\ln g_L.\end{aligned}$$

This is almost identical to the regression run by Frank R. Lichtenberg (1993) on the same sample of countries, for the same period, as Mankiw, Romer and Weil. Lichtenberg reports that inclusion of the R&D variable reduces the estimated sum of capital coefficients $\alpha+\beta$ by between 0.02 and 0.1.[26]

B. Convergence Rates

In neoclassical growth theory with only one kind of capital, the rate of convergence is:

$$(1-\alpha)\left(\delta+g_L+g\right) \tag{14}$$

If our model implied the same formula, then our argument to the effect that α is lower than implied by neoclassical theory would also imply that the rate of convergence was correspondingly higher. Since the neoclassical estimate of $\alpha \simeq 0.7$ is consistent with the evidence of a convergence rate around 0.02, this would imply that our model overpredicted the rate of convergence.

In truth, the model predicts that, given α, the rate of convergence should be strictly less than predicted by neoclassical theory. We measure the rate of convergence in our two-dimensional system by the absolute value of the dominant root of the linearized system.[27] As shown in Appendix A, the system has two real roots μ_1 and μ_2, with:

$$\mu_1 < -\left(1-\alpha\right)\left(\delta+g_L+g\right) < \mu_2 < 0. \tag{15}$$

Thus the absolute value $|\mu_2|$ of the dominant root, which is associated with the turnpike vector DD in Figure 2, is smaller than the convergence rate of the corresponding neoclassical model. The intuitive reason why convergence is slower than in the neoclassical model is that as the economy approaches its steady state from below, the diminishing marginal productivity of capital, which slows down the investment rate as in neoclassical theory, will tend to be offset by the simultaneous rise in the country's relative productivity.

To explore by how much the rate of convergence is affected by the endogeneity of relative productivity, I calibrate the model, using the following values:

long-run growth rate of output per person:	g	0.02
rate of depreciation:	δ	0.03
rate of population growth:	g_L	0.015
steady-state rate of creative destruction:	λn	0.036

The first three values are chosen to represent the US economy over the 1960-85 period, the sample period of Mankiw, Romer and Weil. The rate of creative destruction is chosen from the evidence of Ricardo J. Caballero and Adam B. Jaffe (1993) who estimate that the average US company that does not innovate loses value at a 3.6% annual rate.

As explained in Appendix A, the rate of convergence can be calculated from these values, the capital coefficient α, and the elasticity of the equilibrium R&D function $\widetilde{n}$ with respect to k. This elasticity depends on the R&D subsidy ψ and the number of workers per product ℓ. In addition, it would be affected by any subsidies to (or taxes on) capital accumulation, a factor that I omit for simplicity. The elasticity would also be affected by specifying a realistic degree of decreasing returns to the R&D production function.[28] Rather than calibrate problematical parameters or introduce extra complications I note that the elasticity of $\widetilde{n}$ must be at least as great as the elasticity of the $\widetilde{\pi}$ function with respect to k, namely α.[29] Furthermore, as shown in Appendix A, the absolute size $|\mu_2|$ of the dominant root is decreasing with respect to the elasticity of $\widetilde{n}$.

Accordingly, Figure 4 below plots the convergence rate of the neoclassical model (14) and of the Schumpeterian model ($|\mu_2|$), for different values of α, under the assumption that the elasticity of $\widetilde{n}$ equals its lower bound α. The difference between the two curves would be even larger for any other value of this elasticity.

Figure 4 here

Figure 4 shows by how much the convergence rate falls below that predicted by the neoclassical model, for any given capital coefficient α. Equivalently, it shows by how much the estimated capital coefficient falls below that estimated by the neoclassical model given any convergence rate. The size of the difference decreases as α increases. When the rate of convergence is at its commonly estimated value of .02, the neoclassical model would imply $\alpha = 0.7$ whereas the Schumpeterian model would imply at most $\alpha = 0.6$. These findings are roughly consistent with the bias of up to 0.1 implied by Lichtenberg's cross-country regression.

Thus the interpretation that the Schumpeterian model places on evidence coming from cross-country regressions is compatible with the interpretation it places on estimated rates of convergence. In both cases the model implies a lower coefficient of capital in the aggregate production function than does the neoclassical model. This finding illustrates a sharp distinction between the Schumpeterian version of endogenous growth theory and the AK version. Neoclassical theorists correctly interpret a finding of $\alpha < 1$ as evidence against AK theory, and in favor of neoclassical theory. But, as we have just shown, a small value of α is actually more supportive of Schumpeterian theory than of neoclassical theory.

VIII. Conclusions

In summary, a simple open-economy extension of the Schumpeterian endogenous growth model with technology transfer is capable of accounting for the cross-country evidence that has been adduced by detractors of endogenous growth. The model implies that countries with positive R&D levels will converge to parallel growth paths, with the same positive growth rate, while other countries will stagnate. The common long-run growth rate among growing economies depends on the same parameters that would have influenced growth according to the closed-economy Schumpeterian model. In particular, it will be raised by an increase in the investment rate or the R&D-subsidy rate in any R&D-producing country, although the effect of a single country's parameters may be very small.

The model also implies that these parameter changes will affect a country's relative level of income per person in the long run, and in the same direction as they influence world growth. It predicts exactly the same equation that Mankiw, Romer and Weil (1992) argue explains over three quarters of the cross-country variation in income per person. However, it implies that their procedure for estimating the contribution of differences in equilibrium capital/labor ratios is likely to be biased upward because of the omission of productivity levels from their equation. It also shows that using the neoclassical model to calculate the capital coefficient from estimated convergence rates yields an upward-biased estimate.

Countries differ in per-capita income not only because of differences in capital stocks but also because of differences in productivity. The Schumpeterian model is broadly consistent with existing cross-country evidence and also accounts for the cross-country differences in productivity levels that recent research has shown to be quite large. Because of this, and because it offers an explanation of the steady-state world growth rate, Schumpeterian growth theory offers a more satisfactory framework than the neoclassical approach, which makes no attempt to explain productivity differences or steady-state growth rates.

References

Aghion, Philippe and Peter Howitt. "A Model of Growth through Creative Destruction." *Econometrica,* March 1992, *60*(2), pp. 323-51.

______. *Endogenous growth theory.* Cambridge, MA: MIT Press, 1998.

Barro, Robert J.; N. Gregory Mankiw and Xavier Sala-i-Martin. "Capital Mobility in Neoclassical Models of Growth." *American Economic Review,* March 1995, *85*(1), pp. 103-15.

Barro, Robert J. and Xavier Sala-i-Martin. *Economic growth.* New York: McGraw-Hill, 1995.

______. "Technological Diffusion, Convergence, and Growth." *Journal of Economic Growth,* March 1997, *2*(1), pp. 1-6.

Basu, Susanto and David N. Weil. "Appropriate Technology and Growth." *Quarterly Journal of Economics,* November 1998, *113*(4), pp. 1025-54.

Caballero, Ricardo J. and Adam B. Jaffe. "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth." in Olivier Jean Blanchard and Stanley Fischer, eds., *NBER macroeconomics annual 1993.* Cambridge, MA: MIT Press, 1993, pp. 15-74.

Chaari, V. V.; Patrick J. Kehoe and Ellen R. McGrattan. "The Poverty of Nations: A Quantitative Investigation." Federal Reserve Bank of Minneapolis Research Department Staff Report No. 204, October 1997.

Coe, David T. and Elhanan Helpman. "International R&D Spillovers." *European Economic Review,* May 1995, *39*(5), pp. 859-97.

Dinopoulos, Elias and Peter Thompson. "Reassessing the Empirical Validity of the Human-Capital Augmented Neoclassical Growth Model." *Journal of Evolutionary Economics,* January 1999, *9*(1), pp. 135-54.

Durlauf, Steven N. and Paul A. Johnson. "Multiple Regimes and Cross-Country Growth Behavior." *Journal of Applied Econometrics,* Oct.-Dec. 1995, *10*(4), pp. 365-84.

Eaton, Jonathan and Samuel Kortum. "Trade in Ideas: Patenting and Productivity in the OECD." *Journal of International Economics,* May 1996, *40*(3-4), pp. 251-78.

Evans, Paul. "Using Cross-Country Variances to Evaluate Growth Theories." *Journal of Economic Dynamics and Control,* June-July 1996, *20*(6-7), pp. 1027-49.

Gantmacher, F. R. *Matrix theory.* Vol. I. New York: Chelsea Publishing Company, 1959.

Grossman, Gene M. and Elhanan Helpman. *Innovation and growth in the global economy.* Cambridge, MA: MIT Press, 1991.

______. "Endogenous Innovation in the Theory of Growth," *Journal of Economic Perspectives,* Winter 1994, *8*(1), pp. 23-44.

Hall, Robert E. *Booms and recessions in a noisy economy.* New Haven, CT: Yale University Press, 1991.

Hall, Robert E. and Charles I. Jones. "Why Do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics,* February 1999, *114*(1), pp. 83-116.

Helliwell, John F. and Alan Chung. "Macroeconomic Convergence: International Transmission of Growth and Technical Progress." in Peter Hooper and J. David Richardson, eds., *International economic transactions: Issues in measurement and empirical research.* NBER Studies in Income and Wealth, Vol. 55. Chicago: The University of Chicago Press, 1991, pp. 388-436.

Howitt, Peter. "Steady Endogenous Growth with Population and R&D Inputs Growing." *Journal of Political Economy,* August 1999, *107*(5), pp. 715-30.

Howitt, Peter and Philippe Aghion. "Capital Accumulation and Innovation as Complementary Factors in Long-Run Growth." *Journal of Economic Growth,* June 1998, *3*(2), pp. 111-130.

Islam, Nazrul. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics,* November 1995, *110*(4), pp. 1127-70.

Jones, Charles I. "R&D-Based Models of Economic Growth." *Journal of Political Economy,* August 1995, *103*(4), pp. 759-84.

Jorgenson, Dale W. *Productivity.* Cambridge, MA: MIT Press, 1995.

Jovanovic, Boyan and Yaw Nyarko. "Learning By Doing and the Choice of Technology." *Econometrica,* November 1996, *64*(6), pp. 1299-1310.

Keller, Wolfgang. "How Trade Patterns and Technology Flows Affect Productivity Growth." Working paper, University of Texas, February 1999.

Klenow, Peter J. and Andrés Rodríguez-Clare. "The Neoclassical Revival in Growth Economics: Has it Gone too Far?" in Ben Bernanke and Julio Rotemberg, eds., *NBER macroeconomics annual1997.* Cambridge, MA: MIT Press, 1997, pp. 73-103.

Kortum, Samuel S. "Equilibrium R&D and the Patent-R&D Ratio: U.S. Evidence." *American Economic Review,* May 1993, (*Papers and Proceedings*) *83*(2), pp. 450-57.

______. "Research, Patenting, and Technological Change." *Econometrica,* November 1997, *65*(6), pp. 1389-1419.

Lichtenberg, Frank R. "R&D Investment and International Productivity Differences." in H. Giersch, ed., *Economic growth in the world economy.* Tuebingen, Germany: J.C.B. Mohr, 1993, pp. 89-110.

Mankiw, N. Gregory; David Romer and David N. Weil. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics,* May 1992, *107*(2), pp. 407-37.

McGrattan, Ellen R. and James A. Schmitz, Jr. "Explaining Cross-Country Income Differences." Federal Reserve Bank of Minneapolis Research Department Staff Report No. 250, July 1998.

Parente, Stephen L. and Edward C. Prescott. "Technology Adoption and Growth." *Journal of Political Economy,* April 1994, *102*(2), pp. 298-321.

Quah, Danny T. "Convergence Empirics Across Economies with (Some) Capital Mobility." *Journal of Economic Growth,* March 1996, *1*(1), pp. 95-124.

Trefler, Daniel. "International Factor Price Differences: Leontief Was Right!" *Journal of Political Economy,* December 1993, *101*(6), pp. 961-87.

______. "The Case of Missing Trade and Other Mysteries." *American Economic Review,* December 1995, *85*(5), pp. 1029-46.

Young, Alwyn. "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." *Quarterly Journal of Economics,* August 1995, *110*(3), pp. 641-80.

Notes

*Department of Economics, The Ohio State University, 410 Arps Hall, 1945 North High St., Columbus OH 43210-1172. This paper originated in work with Philippe Aghion, the results of which are contained in our recent book (Aghion and Howitt, 1998). I have benefitted from comments by fellow members of the Economic Growth and Policy Program of the Canadian Institute for Advanced Research, seminar participants at The Ohio State University, Wolfgang Keller, and two referees. Jo Ducey and Young-Kyu Moh provided valuable assistance.

Notes

[1]Since Evans's work is restricted to OECD countries, it leaves open the possibility of "club-convergence." That is, convergence may take place only within a select group of nations, as argued by Stephen N. Durlauf and Paul A. Johnson (1995), Danny T. Quah (1996), and others.

[2]Aghion and Howitt (1998, ch.12) argue that the Schumpeterian branch of endogenous growth theory is also consistent with the evidence of Charles I. Jones (1995) on the falling productivity of R&D and with the growth-accounting exercises of Dale W. Jorgenson (1995) and Alwyn Young (1995) that report a relatively small contribution of productivity to economic growth.

[3]For example, John F. Helliwell and Alan Chung (1991), Stephen L. Parente and Edward C. Prescott (1994), Jonathan Eaton and Samuel Kortum (1996), Barro and Sala-i-Martin (1997), and Susanto Basu and Weil (1998). What distinguishes the present analysis from these papers is that it presents an integrated analysis of capital accumulation and endogenous innovation.

[4]But without the international trade in goods that Coe and Helpman argue is the vehicle for such transfers. Wolfgang Keller (1999) shows that the spillovers depend not only on the size of trade and on the identities of a country's trading partners but also on the industrial pattern of trade flows.

[5]The analysis does not attempt to explain why countries have different investment rates or operate at different efficiency levels, as do Hall and Jones (1999) and V. V. Chaari, Patrick J. Kehoe and Ellen R. McGrattan (1997). Instead it takes each country's investment rate as given and assumes that all countries are equally efficient at using any given technology.

[6]Howitt (1999) derives a closed-economy model with the same basic structure but in which

the horizontal innovations creating new products are motivated by the same profit-seeking objectives as vertical quality-improving innovations.

[7]If there were no population growth, ℓ would fall to zero, and the research-arbitrage equation (7) below would not yield a positive solution to the R&D intensity n_t. In this case the country's growth rate would fall to zero, as discussed in Section V below in connection with club convergence.

[8]The prediction that more populous areas will have proportionately more product variety, all else equal, is common to all models of endogenous product variety. Although this prediction accords with common observation (see for example Hall, 1991, p.7), I know of no test of it using cross-country data on population and product diversity.

[9]An alternative explanation for the lack of an observed scale effects across countries would be that entrepreneurs in even small countries can sell their innovations in a global market. This channel has been ruled out in the present analysis by the assumption of no trade in intermediate goods. Thus an innovation can be marketed only in the country of origin.

[10]Under the Cobb-Douglas technology (2) this has no substantive implications.

[11]No innovations are done by incumbents because of the Arrow- or replacement-effect. (See Aghion and Howitt, 1992).

[12]From (3), the definition of A_t and the adding-up condition, $K_t = \int_0^{N_t} A_t(i)\, x_t di = N_t A_t x_t$. Equation (4) follows from this by the definitions of k_t and ℓ.

[13]Thus when sector i innovates, the proportional increase in $A_t(i)$ will depend on how long it has been since the last innovation in sector i. The alternative assumption, used by Aghion and Howitt (1992) and Gene M. Grossman and Helpman (1991), to the effect that the proportional increase in $A_t(i)$ is a fixed constant, neglects the effect of spillovers coming from innovations in other sectors on the quality of an innovation.

[14]Thus the model embodies the "diminishing opportunities" hypothesis of Kortum (1997). As explained in Howitt (1999), the model is also consistent with Kortum's observation of a declining rate of patenting per R&D scientist/engineer, because we may interpret the increase in scientists and engineers as an increase in the (human) capital input to R&D.

[15]This formulation assumes that the previous incumbent is unable to re-enter once it stops producing. That is why a successful innovator can ignore potential competition from previous innovators in the same product. Howitt and Aghion (1998, Appendix) show that the alternative case in which the previous incumbent is free to reenter produces the same steady-state comparative-statics results in a related closed-economy model.

[16]I assume all countries share the same depreciation rate δ, production kernel f and imitation rate ξ.

[17]Imitations do not affect the evolution of the average A_t, because they result in new sectors whose productivity parameters are drawn at random from the distribution of existing

parameters in the country.

[18]There is no scale effect of the level of population on y. There might or might not be an effect of the rate of population growth g_L, whose effect on y is ambiguous in sign because its effect on the country's relative productivity level is ambiguous in sign.

[19]Thus a country can be stuck in a no-growth situation even though there is no real advantage to backwardness of the sort that Boyan Jovanovic and Yaw Nyarko (1996) show can arise from the dynamics of learning by doing under different technological paradigms.

[20]Club convergence is not contradicted by Evans's (1996) demonstration of σ-convergence, which is restricted to OECD countries, all of whom perform a substantial amount of R&D.

[21]The marginal contribution (σ_j/N_{jt}) has been deflated by the number of products in that country, rather than by the number in the world, in order to avoid a technical problem common to all models of technology transfer with convergence. That is, deflating by the number in the world would lead to a degenerate steady state in which the only country with a measurable effect on world technology is the one with the fastest population growth, since the fraction of world R&D performed in that country will approach unity in the very long run. Thus the present model's steady state depicts a medium-long run in which no country's population growth has yet overwhelmed the rest of the world.

[22]Kortum (1997, esp. pp.1400-1) provides an alternative derivation of the relationship between R&D intensity and productivity growth, which is not consistent with the proportional form of (S).

[23]Equation (WG) implies that increasing the number of countries would increase the world growth rate. This prediction depends critically however on the simplifying assumption discussed in footnote 21. Even if we were to relax this assumption the world growth rate would not be affected by the size of world population.

[24]The identifying assumption has also been challenged by such writers as Nazrul Islam (1995) and Hall and Jones (1999). Others have challenged the Mankiw-Romer-Weil results on the grounds that growth and investment rates are simultaneously determined and that alternative measures of human capital produce different results. See, for example, Peter J. Klenow and Andrés Rodríguez-Clare (1997) or Elias Dinopoulos and Peter Thompson (1999). For a recent survey of the issues involved with the Mankiw-Romer-Weil results, see McGrattan and James A. Schmitz Jr. (1998).

[25]Grossman and Helpman (1994) observe that investment rates are positively correlated with productivity growth rates, which biasses the estimated capital coefficients in the same direction.

[26]Lichtenberg's equation contains the same variables except for g_L, which is not included as a separate regressor. He finds, using five alternative definitions of R&D, that inclusion of an R&D variable reduces the estimated value of $\alpha + \beta$ by an amount that varies from 0.017 to 0.090. (See his Table 3.) He also reports two dynamic regressions (with the 1960 income

level included as an additional regressor), in which the inclusion of the R&D variable reduces $\alpha + \beta$ by 0.095 and 0.097. (His Table 4.)

[27]This is the procedure implicit in neoclassical theory, in which the dynamics can be represented as the limiting case of a system:

$$\begin{aligned} \dot{k}_t &= sk_t^{\alpha} - \left[\delta + g_L + \Lambda\left(a_t^{-1} - 1\right)\right] k_t \\ \dot{a}_t &= \Lambda\left(1 - a_t\right) - a_t g \end{aligned}$$

in which the speed of automatic technology transfer $\Lambda > 0$ is arbitrarily large. The dominant root of the linearized version of this system is minus the neoclassical convergence rate (14), the other root being arbitrarily large.

[28]That is, I assume an elasticity of innovations with respect to R&D equal to unity, whereas Kortum (1993) estimates values of this elasticity between 0.1 and 0.6 at the industry level.

[29]From (5) and (7):

$$\widetilde{n}\left(k;\theta\right) = \frac{\widetilde{\pi}\left(k\right)\ell}{1-\psi} - \frac{\alpha f'\left(k\right) - \delta}{\lambda}.$$

Since $\alpha f' - \delta = r > 0$ and $f'' < 0$, it follows from this and (6) that:

$$\frac{k}{\widetilde{n}}\frac{\partial \widetilde{n}}{\partial k} \geq \frac{k}{\widetilde{\pi}}\frac{\partial \widetilde{\pi}}{\partial k} = \alpha.$$

Appendix A

This appendix derives the roots of the local approximation to (9) and (10) of Section III. Linearizing around the steady state $\left(\hat{a}, \hat{k}\right)$ yields:

$$\begin{bmatrix} \dot{a} \\ \dot{k} \end{bmatrix} = M \begin{bmatrix} a - \hat{a} \\ k - \hat{k} \end{bmatrix}$$

with Jacobian matrix:

$$M = \begin{bmatrix} -\lambda\widetilde{n} - g & \lambda\frac{\partial\widetilde{n}}{\partial k}\left(1-a\right) \\ \lambda\widetilde{n}k/a^2 & \alpha s k^{\alpha-1} - \delta - g_L - g - \lambda\frac{\partial\widetilde{n}}{\partial k}\left(a^{-1}-1\right)k \end{bmatrix},$$

which can be rewritten, using the steady-state equations (A) and (K), as:

$$M = \begin{bmatrix} -z_0 & z_1 \\ \vartheta z_0 & -z_2 - \vartheta z_1 \end{bmatrix},$$

where:

$$\text{(A1)} \qquad \begin{cases} z_0 & = \lambda\widetilde{n} + g > 0, \\ z_1 & = \lambda\frac{\partial\widetilde{n}}{\partial k}\left(1-a\right) > 0, \\ z_2 & = \left(1-\alpha\right)\left(\delta + g_L + g\right) > 0, \text{ and} \\ \vartheta & = k/a. \end{cases}$$

Note that, by (A):

$$\vartheta z_1 = \eta g,$$

where η is the elasticity $\frac{k}{\widetilde{n}}\frac{\partial\widetilde{n}}{\partial k}$. Thus the roots of M are:

$$\mu_1 = (1/2)\left(\text{tr } - \sqrt{\text{tr}^2 - 4\det}\right) \text{ and } \mu_2 = (1/2)\left(\text{tr } + \sqrt{\text{tr}^2 - 4\det}\right),$$

where the values of:

$$\text{(A2)} \qquad \begin{cases} \text{tr} = -\left(z_0 + z_2 + \eta g\right) \\ \det = z_0 z_2 \end{cases}$$

are determined according to (A1) by $\alpha, g, \delta, g_L, \lambda\widetilde{n}$ and η. It follows from the sign restrictions in (A1) that μ_1 and μ_2 are both real and that $\mu_1 < \mu_2 < 0$.

If η were zero then (A2) would solve for $\mu_1 = -\max\{z_0, z_2\}$ and $\mu_2 = -\min\{z_0, z_2\}$, which would imply that $\mu_1 \leq -z_2 \leq \mu_2$. Since μ_1 is decreasing and μ_2 is increasing with respect to η, holding constant α, g, δ, g_L and λn, and since η must be strictly positive (see footnote 29) therefore: $\mu_1 < -z_2 < \mu_2 < 0$. Since z_2 is the neoclassical speed of convergence (14) this establishes the inequalities (15) in the text.

Appendix B

This appendix shows that the system composed of equations (11) and (12) for all j is locally stable. The Jacobian of its linear approximation can be partitioned into four square matrices:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

The steady-state conditions (A) and (K) for each country and the spillover equation (S) can be used to express these matrices as:

$$\begin{aligned}
A &= -\mathrm{diag}\left(\lambda_j n_j + g\right) = -\mathrm{diag}\left(z_{0j}\right), \\
B &= \mathrm{diag}\left(\left(1-a_j\right)\lambda_j \frac{\partial \widetilde{n}_j}{\partial k_j}\right) - v \cdot u' = \mathrm{diag}\left(z_{1j}\right) - v \cdot u', \\
C &= \mathrm{diag}\left(\lambda_j n_j k_j / a_j^2\right) = \mathrm{diag}\left(\vartheta_j z_{0j}\right), \text{ and} \\
D &= -\mathrm{diag}\left(\left(1-\alpha\right)\left(\delta + g_{Lj} + g\right)\right) - \mathrm{diag}\left(\lambda_j \frac{\partial \widetilde{n}_j}{\partial k_j}\left(a_j^{-1} - 1\right) k_j\right) = -\mathrm{diag}\left(z_{2j}\right) - \mathrm{diag}\left(\vartheta_j z_{1j}\right);
\end{aligned}$$

where:

$$v = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \text{ and } u = \begin{bmatrix} \sigma_1 \lambda_1 \frac{\partial \widetilde{n}_1}{\partial k_1} \\ \vdots \\ \sigma_m \lambda_m \frac{\partial \widetilde{n}_m}{\partial k_m} \end{bmatrix},$$

and where z_{0j}, z_{1j}, z_{2j} and ϑ_j are country-specific versions of the variables defined in (A1).

We want to show that the characteristic roots of M all have negative real parts. These roots are the solutions to the characteristic equation:

$$P\left(\mu\right) \equiv \left|M - \mu I\right| = 0.$$

To this end, consider any such root μ. If μ is also a characteristic root of the matrix A or the matrix D, then it is real and negative, since A and D are negative diagonal matrices.

So suppose μ is not a root of A or D. Then by the rules of partitioned determinants (F. R. Gantmacher, 1959, pp.45-6) we can write the characteristic function P as:

$$P\left(\mu\right) = \left|A - \mu I\right| \cdot \left|D - \mu I - C\left(A - \mu I\right)^{-1} B\right|.$$

Hence:

$$P(\mu) = \left|(A - \mu I)(D - \mu I) - (A - \mu I) C (A - \mu I)^{-1} B\right|.$$

Since A and C are diagonal:

$$P(\mu) = |(A - \mu I)(D - \mu I) - CB|.$$

Expanding the product CB and gathering terms yields:

(B1)
$$P(\mu) = |\operatorname{diag}(Q_j(\mu)) + \tilde{v} \cdot u'|,$$

where:

$$\tilde{v} = \begin{bmatrix} \vartheta_1 z_{01} a_1 \\ \vdots \\ \vartheta_m z_{0m} a_m \end{bmatrix},$$

and each $Q_j(\cdot)$ is the characteristic function of the country-specific matrix:

$$M_j = \begin{bmatrix} -z_{0j} & z_{1j} \\ \vartheta_j z_{0j} & -z_{2j} - \vartheta_j z_{1j} \end{bmatrix},$$

both of whose roots we have seen (in Appendix A) are real and negative. By expanding the determinant in (B1) we arrive at:

(B2)
$$P(\mu) = \prod_{j=1}^{m} Q_j(\mu) + \sum_{j=1}^{m} \tilde{v}_j u_j \prod_{\substack{i=1 \\ i \neq j}}^{m} Q_i(\mu).$$

If μ is a root of one of the Q_j's, then, as we have seen, it must be real and negative. So suppose it is not a root of any Q_j. Then, from (B2):

(B3)
$$1 + \sum_{j=1}^{m} \frac{\tilde{v}_j u_j}{Q_j(\mu)} = 0.$$

We need only show that all the roots of (B3) have negative real parts.

Since each $Q_j(\mu)$ has a positive quadratic coefficient and both roots of each Q_j are real and negative, therefore:

(B4)
$$Q_j(\mu) > 0 \text{ for all } j = 1, ..., m \text{ if } \mu \geq 0.$$

Since each $\tilde{v}_j u_j$ is positive, (B4) implies that if μ is a real root of (B3) it must be negative. So we need only show that the real part of any complex root of (B3) is negative.

Let $\mu = \zeta + \omega i$ be a complex root of (B3). Since its complex conjugate is also a root, assume with no loss of generality that $\omega > 0$. For each j let $Q_j(\mu) = \tau_j + \chi_j i$ and $1/Q_j(\mu) = \rho_j + \varepsilon_j i$, and let μ_{1j} and μ_{2j} be the two (negative) roots of Q_j. Then:

$$\rho_j = \frac{\tau_j}{\tau_j^2 + \chi_j^2} \text{ and } \varepsilon_j = -\frac{\chi_j}{\tau_j^2 + \chi_j^2} \text{ for all } j = 1, ..., m.$$

Also, since each $Q_j(\mu) = \left(\mu_{1j} - \mu\right)\left(\mu_{2j} - \mu\right)$, therefore:

$$\chi_j = \omega\left(-\mu_{1j} - \mu_{2j} + 2\zeta\right) \text{ for all } j = 1, ..., m.$$

It follows that the real part ζ of μ must be negative or else the imaginary part $\tilde{v}_j u_j \varepsilon_j = \tilde{v}_j u_j \omega\left(\mu_{1j} + \mu_{2j} - 2\zeta\right) / \left(\tau_j^2 + \chi_j^2\right)$ of each expression $\tilde{v}_j u_j / Q_j(\mu)$ would be negative, so that μ would not solve (B3).$\|$

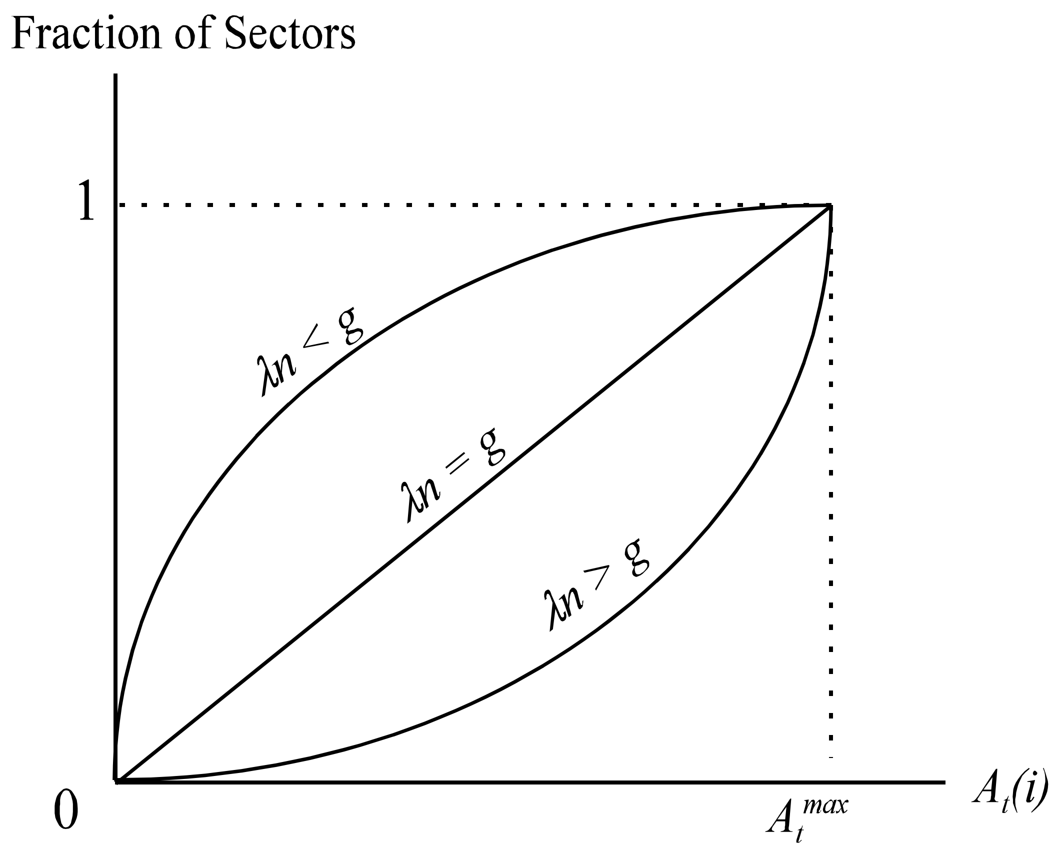


FIGURE 1: The cross-sectional distribution of productivity parameters within a country

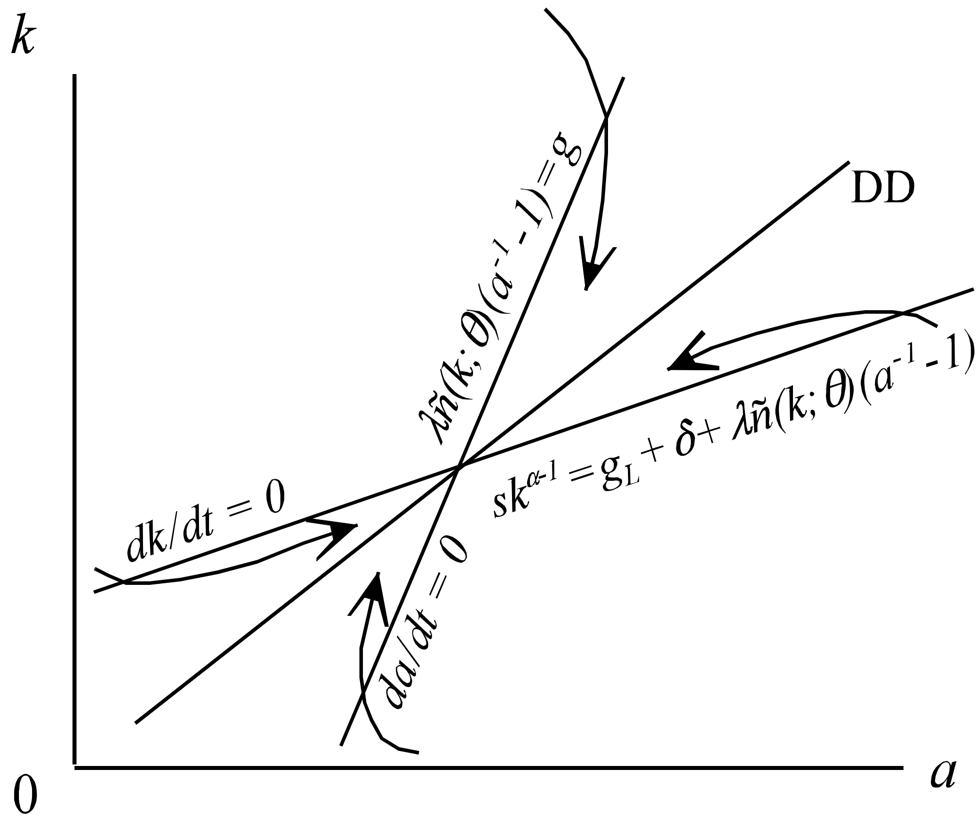
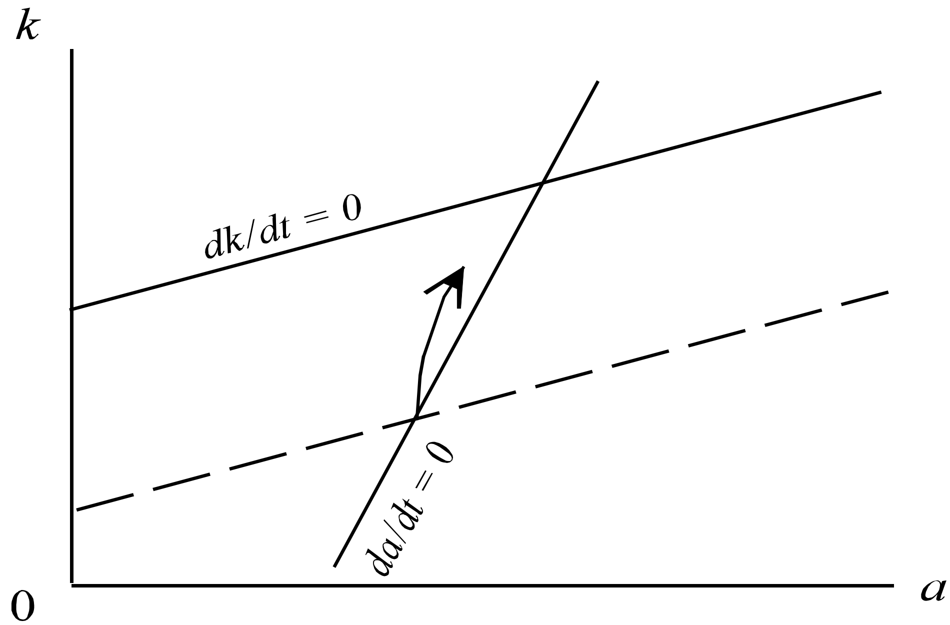
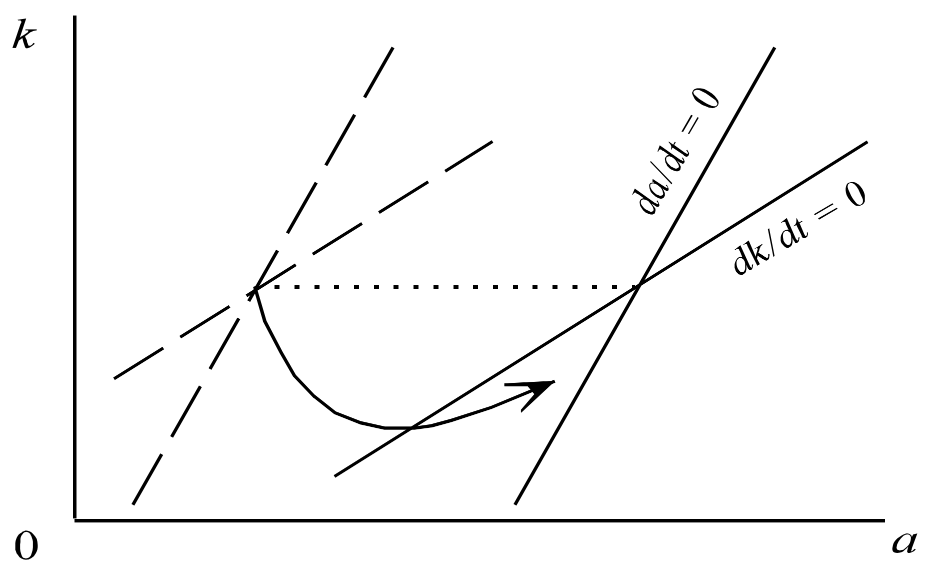


FIGURE 2: The phase portrait of a country's relative productivity (a) and capital stock per effective worker (k)



(a)



(b)

FIGURE 3: Reactions to (a) an increase in the investment rate and (b) an increase in R&D productivity or in the R&D subsidy rate

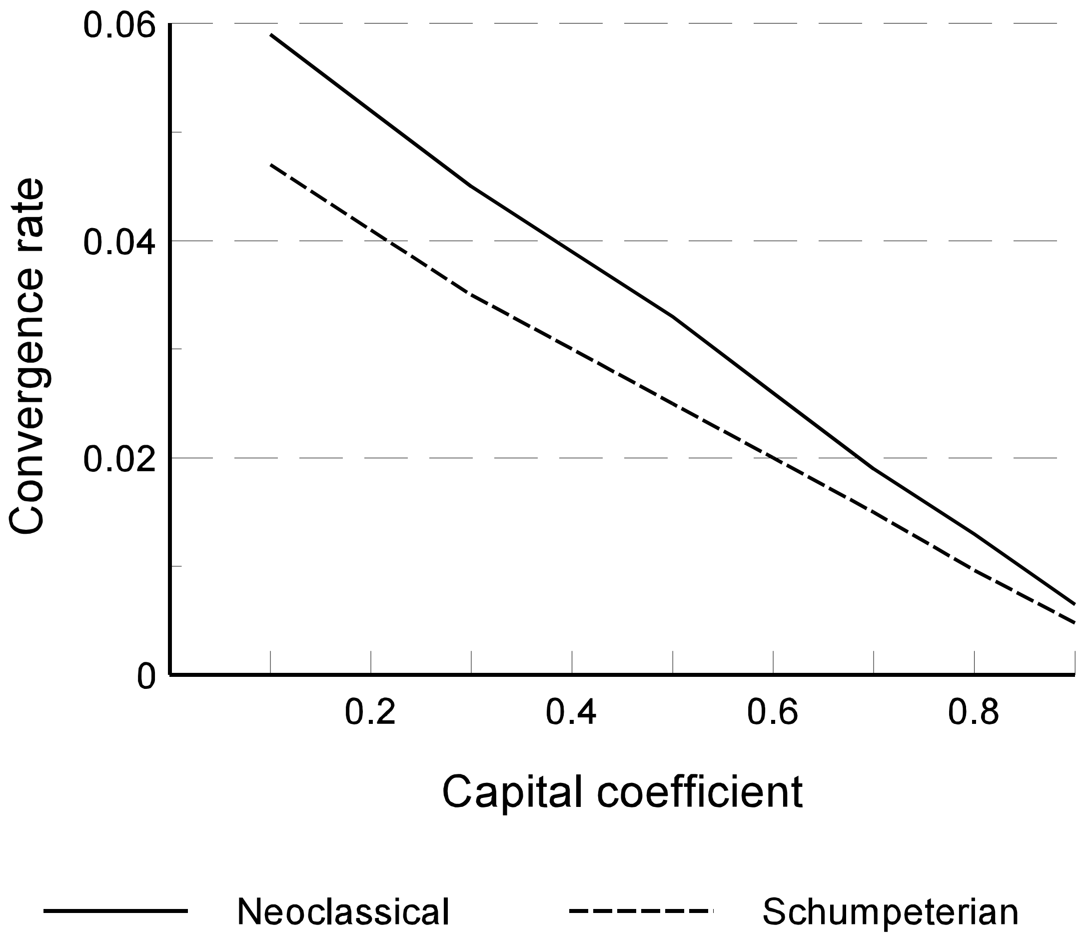


FIGURE 4: The relationship between capital's coefficient and the convergence rate