IUI

THE INDUSTRIAL INSTITUTE FOR ECONOMIC AND SOCIAL RESEARCH

WORKING PAPER No. 475, 1996

LECTURE NOTES ON KNOWLEDGE AND HUMAN CAPITAL IN THE NEW GROWTH THEORY

BY PAUL SEGERSTROM

Lecture Notes on Knowledge and Human Capital in the New Growth Theory

Paul Segerstrom
Michigan State University
December 7, 1996

Paper presented at **The Astra-Ericsson Seminar Series on Human Capital and Economic Growth,** December 17 and 19, 1996.

1 Introduction

Long-run economic growth—by which I mean a continuing rise in per-capita real GDP over a long period of time—is one of the most important topics in economics. The newspapers are filled with accounts of monthly changes in economic activities. But these short-run changes have a relatively minor impact on economic well-being. Why industrial production rose or fell a few percent over the last three months can be an intriguing question. Even more significant, however, is why standards of living are so much higher today than 20, 50 or 100 years ago. Seemingly modest growth rates, if sustained over a long period of time, have an enormous effect on living standards.

Take the United States as an example. Table 1 illustrates the average annual per-capita real GDP growth rate over the last 200 years. Although the growth numbers in the table may seem small, being between 1 and 2 percent, the fact that this modest growth was sustained over a period of roughly two hundred years has had an enormous effect on peoples' lives. Consider the following statistics: In 1800, 90 percent of the labor force worked on farms

Table 1: Average Annual Real Per-Capita GDP Growth Rates in the United States

Time Period	Growth Rate
1800-1855	1.1%
1855-1900	1.6%
1900-1950	1.7%
1950-1995	1.9%

and barely produced enough food to feed the country. Today 3 percent of Americans works on farms and there is a surplus of food (the government pays farmers to take land out of production). In the mid 1800's, the typical diet consisted of potatoes, lard, cornmeal and salt pork (which is not very inspiring to eat day in and day out). In the rural U.S. in the 1800's, houses were tiny and crudely built, there were no glass windows, no lighting except the fireplace, and no indoor plumbing. Urban housing in the 1800's was not much better. For example, in the 1860's, in New York City, it was typical for 6 people to live in a 10 by 12 foot room. In Boston, there was 1 bathtub for every 50 people, no homes had electricity, and less than 2 percent of the population had indoor toilets. Also life expectancy in the 1800's was low, in part because diseases like smallpox, diphtheria, typhoid fever and whooping cough were still common and in part because women often died in childbirth. Whereas the average life expectancy today is around 75 years, as late as 1900, the average life expectancy was only 47 years.[1]

The truth is that life just 100 years ago was brutal, filled with hard work and little food or health care. Lifetimes were short and infant mortality rates were high. What has happened since then to change peoples lives is long-run economic growth—a continuing rise in per capita real GDP over a long period of time. Economic growth has lifted billions of people out of extreme proverty. Perhaps the most important issue that a theory of economic growth should address is, why has economic growth occurred?

This brings me to a second important topic: not all countries have experienced economic growth to the same extent, resulting in large differences in living standards across countries (see Table 2). As is illustrated in Table 2, compared to the 1985 GDP per adult for the

Table 2: 1985 Real GDP Per Adult For Selected Countries (in U.S. Dollars)

Country	GDP/Adult
Kuwait	25,635
United States	18,988
Sweden	15,237
Burundi	663
Chad	462
Zaire	412

United States of $18,988, Chad's GDP per adult was $462 and Zaire's was $412.[2] Although this difference of a factor of 40 in living standards is too large to be believed, there is clearly

[1]See Stockman [1996] and Baumol and Blinder [1988].

[2]See Mankiw, Romer and Weil [1992].

Table 3: Annual Rates of Per Capita Real GDP Growth in Various Countries

Country	Growth Rate, 1900-1994	Country	Growth Rate, 1950-1994
Japan	3.2%	Taiwan	6.1%
Finland	2.6%	Japan	6.0%
Canada	2.3%	South Korea	5.7%
Switzerland	2.0%	Germany	3.6%
Australia	1.5%	United States	1.9%
United Kingdom	1.4%	India	1.7%
Argentina	1.1%	Chile	1.0%
Bangladesh	0.1%	Bangladesh	0.3%

considerable diversity across countries in measured per capita income levels. Furthermore, as is illustrated in Table 3, there is considerable diversity in real per capita GDP growth rates across countries, even over sustained periods.[3] From 1950-1994, we observe, for example: Bangladesh, 0.3% per year; India, 1.7%; Germany, 3.6%; Japan., 6.0%. The economies of some countries, such as Hong Kong, Japan, Singapore, South Korea and Taiwan have grown much faster than that of the United States in recent years, while other countries, such as Zambia, Madagascar, and Bangladesh, have grown much more slowly. To obtain from the above-mentioned growth rates the number of years it takes for incomes to double, divide the natural log of 2 by a country's growth rate.[4] Then these numbers suggest that Indian incomes will double every 40 years; and Korean incomes will double every 12 years. An Indian will, on average, be twice as well off as his grandfather; a Korean 17 times. Lucas [1988] writes

> "I do not see how one can look at figures like these without seeing them as representing *possibilities*. Is there some action a government of India could take that would lead the Indian economy to grow like [South Korea's]? If so, what,

[3]See Stockman [1996].

[4]$Y(t) = Y(0)e^{gt} = 2Y(0)$ implies that $t = (\ln 2)/g$.

exactly?... The consequences for human welfare involved in questions like these are simply staggering: Once one starts to think about them, it is hard to think about anything else."

In recent years, considerable progress has been made at understanding the determinants of long-run economic growth, although of course, there is still a long way to go. Many models have been developed to shed light on why growth occurs and why countries have grown at different rates for sustained periods of time. Also new data has become available and economists have devoted a great deal of effort at testing the implications of the new theories of economic growth. The purpose of this paper is to provide an introduction to the rapidly expanding "new growth theory" literature.

Given time and space constraints, I will focus in this paper on presenting one branch of the literature that I believe is particularly promising, namely, models where the engine driving economic growth is knowledge creation or technological change. The inspiration behind this branch of the new growth literature comes from the writings of Schumpeter [1942], who argued that improvements in technology have been the real force behind perpetually rising standards of living. This view is now shared by many (but certainly not all) economists.[5] For example, Grossman and Helpman [1994] write,

> "As yet, no empirical study proves that technology has been the engine of modern-day growth. Still, we ask the reader to ponder the following: What would the last century's growth performance have been like without the invention and refinement of methods for generating electricity and using radio waves to transmit sound, without Bessemer's discovery of a new technique for refining iron, and without the design and development of products like the automobile, the airplane, the transistor, the integrated curcuit, and the computer?"

[5]See, for example, Solow [1994], D. Romer [1996], Grossman and Helpman [1994] and P. Romer [1994]. In the earliest contribution to the new growth literature, Romer [1986], the accumulation of private "knowledge capital" by firms drives economic growth. Romer has since abandoned this approach, arguing that his earlier modeling of knowledge capital as a rival input in production is unsatisfactory and represented a theoretical "sleight of hand" [1994,p.15]. Whereas private capital goods are rival inputs in production, knowledge is an inherently non-rival input which can be used simultaneously by many firms and/or workers.

According to Schumpeter [1942], private profit maximizing firms play a central role in the knowledge creation process. Firms engage in knowledge creation activities (R&D) with the goal of driving rival firms out of business.

> "But in capitalist reality as distinguished from its textbook picture, it is not (price) competition which counts but competition from the new commodity, the new technology, the new source of supply, the new type of organization, ...competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and outputs of the existing firms but at their foundations and their very lives." (p.24)

The expectation of earning monopoly profits in the future from innovating leads firms to incur the upfront costs of R&D today.

> "Was not the observed performance [of technological progress] due to that stream of inventions that revolutionized the techniques of production rather than to the businessman's hunt for profits? The answer is in the negative. The carrying into effect of those technological novelties was the essence of that hunt." (p.110)

Schumpeter also emphasized in his writings that the standard assumption in economic theory that markets are perfectly competitive must be relaxed to understand economic growth:

> "Perfect competition implies free entry into every industry....But perfectly free entry into a *new* field may make it impossible to enter it at all. The introduction of new methods of production and the new commodities is hardly conceivable with perfect competition from the start. And this means that the bulk of what we call economic progress is incompatable with it. As a matter of fact, perfect competition is and always has been temporarily suspended whenever anything new is being introduced." (p.45)

The earliest knowledge creation (or Schumpeterian) models were developed by P. Romer [1990], Segerstrom, Anant and Dinopoulos [1990] and Aghion and Howitt [1992]. These models explore the incentives firms have to engage in R&D activities and discover new products or processes (that is to say, develop new ideas). A key feature which distinguishes these models from earlier growth models is that the assumption of perfectly competitive product markets is relaxed. In all three papers, firms that innovate earn monopoly profits, at least temporarily, as a reward for their past R&D efforts. In this paper, two knowledge creation models are examined in detail, an early model by Grossman and Helpman [1991] which synthesizes elements of Segerstrom, et al. [1990] and Aghion and Howitt [1992], as well as a recent model by Segerstrom [1996], which illustrates how the theory has evolved over time. I will show that small differences in assumptions about the returns to investing in R&D by firms generate large differences in equilibrium and welfare implications.

Although knowledge creation models are useful for thinking about why economic growth occurs and how public policies (applied throughout the world) influence the growth process, at least given the current state of theoretical development, these models are not that useful for explaining cross-country differences in income levels and/or growth rates. To better understand the enormous diversity across countries in standards of living, economists have turned to models of human capital accumulation. For example, D. Romer [1996] writes:

> "The main source of differences in standards of living is not different levels of knowledge or technology, but differences in whatever factors allow richer countries to take advantage of advanced technology. Understanding differences in incomes therefore requires understanding the reasons for the differences in these factors." (p.122)

To illustrate the importance of human capital for understanding cross-country differences in real incomes, I present in detail a human capital accumulation model developed by Mankiw, Romer and Weil [1992]. In this model, the engine driving economic growth is exogenous technological change (as in the "old growth theory") and different rates of human and physical capital accumulation across countries are used to explain cross-country differences in standards of living. Mankiw, Romer and Weil test their model empirically and find that the model can account for almost 80 percent of the cross-country variation in income levels. This is no small accomplishment. After presenting the model, I will review the statistical evidence and discuss some of the criticisms that have been raised by other economists.

The rest of this paper is organized as follows: In sections 2 and 3, the knowledge creation

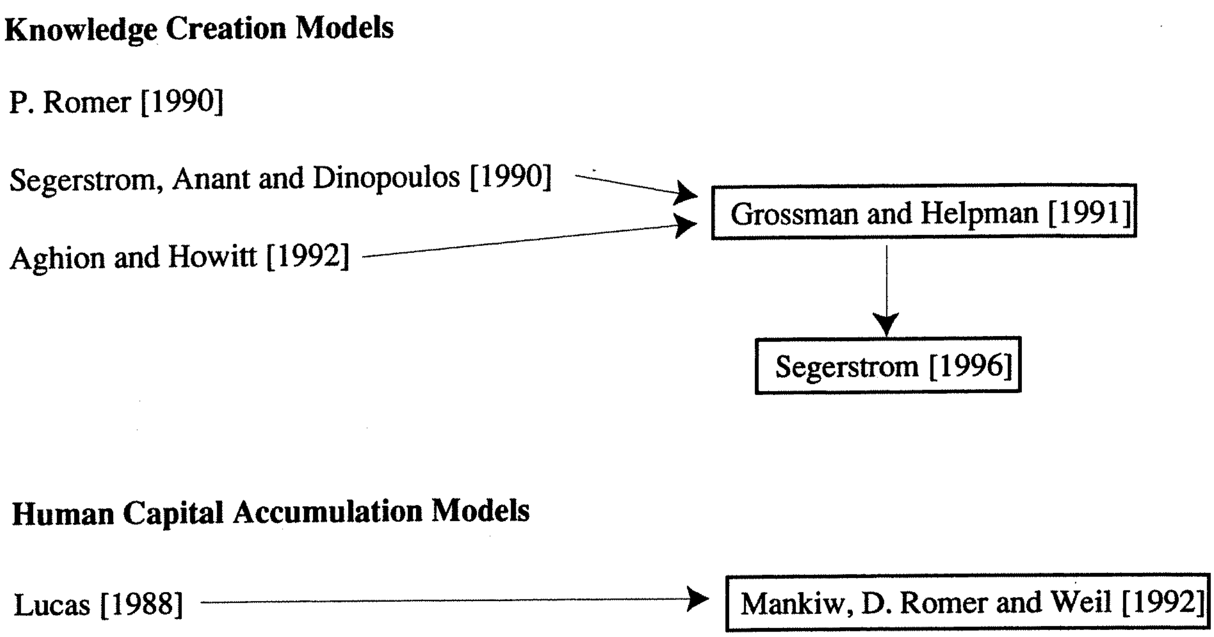

Figure 1: Outline of the Paper

models developed by Grossman and Helpman [1991] and Segerstrom [1996] are presented. Section 4 is devoted to exploring theoretically and empirically the properties of the Mankiw, Romer and Weil [1992] model of human capital accumulation.[6] Some concluding comments are offerred in section 5.

2 Grossman and Helpman [1991]

In this section, the knowledge creation model developed by Grossman and Helpman [1991] is described. This model is often referred to in the literature as a "quality ladders" growth

[6] There is a separate branch of the new growth theory literature which emphasizes human capital accumulation itself as the engine of growth. In Lucas [1988], Jones and Manuelli [1990], and Rebelo [1991], growth in the stock of human capital is the source of per capita income growth. Empirical support for these models is found in Barro and Sala-i-Martin [1995] and Benhabib and Spiegel [1994], where it is documented that schooling is positively correlated with growth of per capita income across countries. However, establishing correlation still leaves open the issue of causation. Is schooling causing growth or is growth causing schooling? In faster growing economies, people have incentives to stay in school longer and acquire more human capital. Bils and Klenow [1996] study this issue and conclude that economic growth is more likely to cause schooling than the other way around.

model, for reasons that will become apparent.

This is a model of an economy with a continuum of industries indexed by $\omega \in [0,1]$. In each industry, firms are distinguished by the quality j of the products they produce. Higher values of j denote higher quality and j is restricted to taking on integer values. At time $t = 0$, the state-of-the-art quality product in each industry is $j = 0$, that is, some firm in each industry knows how to produce a $j = 0$ quality product and no firm knows how to produce any higher quality product. To learn how to produce higher quality products, firms in each industry engage in R&D races.[7] In general, when the state-of-the-art quality in an industry is j, the next winner of a R&D race becomes the sole producer of a $j+1$ quality product. Thus, over time, products improve as innovations push each industry up its "quality ladder."[8]

2.1 Consumer Preferences

All consumers live forever, have identical preferences and maximize discounted utility

$$U \equiv \int_0^\infty e^{-\rho t} \log u(t) dt \tag{1}$$

subject to the usual intertemporal budget constraint. In (1), ρ is the common subjective discount rate, and $\log u(t)$ is the consumer's static utility at time t. The consumer's static utility at time t is in turn given by

$$\log u(t) \equiv \int_0^1 \log \left[\sum_j \lambda^j d(j, \omega, t) \right] d\omega \tag{2}$$

where $d(j, \omega, t)$ denotes the quantity consumed of a product of quality j produced in industry ω at time t, and $\lambda > 1$ represents the extent to which higher quality products improve upon

[7] According to Scherer [1980, p. 409], 59% of the R&D undertaken by firms is aimed at product improvement. The quality ladders growth model can be reformulated as a model of cost reduction without any significant changes.

[8] A large proportion of the scientific research conducted in the OECD countries is financed by private firms. For example, in the United States alone, there are more than 12,000 industrial research labs actively searching for profitable innovations. In Japan, more than 80% of all R&D is financed by private industry. See Rosenberg and Nelson [1993].

lower quality products.

At each point in time t, each consumer i allocates expenditure $c(t)$ to maximize $\log u(t)$ given the prevailing market prices. Solving this optimal control problem yields a unit elastic demand function ($d = c(t)/p$ where d is quantity demanded and p is the relevant market price) for the product in each industry with the lowest quality adjusted price. The quantity demanded for all other products is zero. To break ties, I assume that when quality adjusted prices are the same for two products of different quality, a consumer only buys the higher quality product.

Given this static demand behavior, each consumer i chooses expenditures $c(t)$ over time to maximize U subject to an intertemporal budget constraint. Solving this optimal control problem yields the usual intertemporal optimization condition

$$\frac{\dot{c}(t)}{c(t)} = r(t) - \rho. \tag{3}$$

where $r(t)$ is the market interest rate at time t. In any steady state equilibrium, the left-hand side of (3) is zero, so the market rate of interest must equal the subjective discount rate at each moment in time. Any higher market interest rate induces consumers to save more now and spend more later, resulting in increasing consumer expenditure over time. The steady state level of expenditure c is determined by consumer i's steady state asset holdings.

2.2 Product Markets

In each industry, firms compete in prices. Labor is the only input in production and there are constant returns to scale. One unit of labor is required to produce one unit of ouput, regardless of quality. The labor market is perfectly competitive and for convenience, I normalize the wage of labor to equal one throughout time. Then firms always have constant marginal costs equal to one.

To determine static Nash equilibrium prices and profits, consider any industry $\omega \in [0, 1]$ where there is one quality leader and one follower firm (one step down in the quality ladder). This turns out to be the only type of industry configuration that occurs in equilibrium. With the follower firm charging a price of one, the lowest price it can charge and not lose money, the quality leader earns the profit flow $\pi(p) = (p-1)C/p$ from charging the price p if $p \leq \lambda$, and zero profits otherwise (where $C \equiv \sum_i c$ represents aggregate consumer expenditure). These profits are maximized by choosing the limit price $p = \lambda > 1$. Thus the quality leader earns the profit flow

$$\pi^L \equiv \left(\frac{\lambda - 1}{\lambda}\right) C \tag{4}$$

and none of the other firms in the industry can do better than break even (by selling nothing at all). In this model, new products drive old products from the market unlike in Romer [1990], where new products are no better than old products and old products never become obsolete.[9]

2.3 R&D Races

Labor is the only input used to do R&D in any industry and is perfectly mobile across industries and between production and R&D activities. A firm i that hires ℓ_i units of R&D labor in industry ω at time t is successful in discovering the next higher quality product with instantaneous probability ℓ_i/a, where $a > 0$ is a R&D coefficient. That is, $\ell_i dt/a$ is the probability that the firm will innovate by time $t + dt$ conditional on not having innovated by time t (where dt is an infinitesimal increment of time). Thus, R&D costs are incurred first and profits typically come later in time.[10]

The returns to engaging in R&D races are independently distributed across firms, across industries, and over time. Thus, the industry-wide instantaneous probability of innovative success is simply $I \equiv L_I/a$, where $\sum_i \ell_i = L_I$ is the industry-wide employment of R&D labor. Since this success probability is a linear functions of L_I, each R&D race is characterized by constant returns to R&D.

Let $v(t)$ denote the expected discounted reward for winning a R&D race at time t (or the *value* of being a quality leader). Then at each moment in time, a firm i chooses it's

[9]The fact that quality leaders charge a markup over marginal cost is an essential feature of this model. Quality leaders must earn positive profits if they are to recover their upfront outlays on R&D. Some imperfect competition is necessary to support private investments in new technologies.

[10]An important feature of this R&D technology is that it is available to every firm. Experience in the improvement of a particular product does not provide a lab with future advantages in the improvement of this product. Whatever learning that takes place during the innovation process becomes public.

R&D labor input ℓ_i to maximize its expected profits $\frac{v(t)}{a}\ell_i - \ell_i$. If $v(t) > a$, then $\ell_i = +\infty$ is profit maximizing and if $v(t) < a$, then $\ell_i = 0$ is profit maximizing. Thus it is only profit maximizing for firms to devote a positive (finite) amount of labor to R&D if $v(t) = a$.[11]

The value of a quality leader at time t, $v(t)$, can be determined using the usual arbitrage reasoning. Over a time interval dt, the shareholder receives a dividend $\pi^L(t)\,dt$, and the value of the quality leader appreciates by $\dot{v}(t)\,dt$ in each industry. Because each quality leader is targeted by other firms that conduct R&D to discover the next higher quality product, the shareholder suffers a loss of $v(t)$ if further innovation occurs. This event occurs with probability $I(t)\,dt$, whereas no innovation occurs with probabillity $[1 - I(t)]\,dt$. Efficiency in financial markets requires that the expected rate of return from holding a stock of a quality leader is equal to the riskless rate of return $r(t)\,dt$ that can be obtained through complete diversification: $\frac{\pi^L}{v}dt + \frac{\dot{v}}{v}(1 - I\,dt)\,dt - \left[\frac{v-0}{v}\right] I\,dt = r\,dt$. Taking limits as dt approaches zero, I obtain:

$$v(t) = \frac{\pi^L(t)}{r(t) + I(t) - \frac{\dot{v}(t)}{v(t)}}. \tag{5}$$

The profits earned by each leader π^L are appropriately discounted using the interest rate r and the instantaneous probability I of being driven out of business by further innovation. Also taken into account in (5) is the possibility that these discounted profits grow over time.[12]

The model can be solved for a *symmetric* steady state equilibrium where aggregate consumer expenditure C is constant over time, and industry-level R&D employment L_I does not vary either across industries or over time. When C is a constant, (3) implies that the market interest rate must equal ρ throughout time. Since $v(t) = a$ for all t, (5) implies that in a symmetric steady state equilibrium with positive growth,

$$v \equiv \frac{\left(\frac{\lambda-1}{\lambda}\right)C}{\rho + I} = a. \tag{6}$$

[11] Davidson and Segerstrom [1996] obtain that the profit maximizing innovation rate I is an upward sloping continuous function of the reward for winning v by assuming that there are static decreasing returns to R&D. With constant returns to R&D, as Grossman and Helpman [1991] assume, one does not obtain such a relationship.

[12] In this model, perfect patent protection for innvators is assumed. In Segerstrom [1991] and Davidson and Segerstrom [1996], firms copy other firms' products in equilibrium.

must hold. In (6), the profit flow π^L earned by the winner of a R&D race is discounted using the market interest rate ρ and the instantaneous probability I that the firm will be driven out of business by further innovation.

2.4 The Labor Market

The endowment of labor L in the economy is constant over time. In each industry, C/λ workers are employed in production and $L_I = aI$ workers are employed in R&D. With a measure one of industries, full employment of labor implies that

$$L = \frac{C}{\lambda} + aI. \tag{7}$$

Labor market clearing requires that employment in manufacturing plus employment in R&D equals the available labor supply.

Both the R&D condition (6) and the labor market condition (7) are illustrated in Figure 2. These two steady state conditions have a unique intersection at $I^* > 0$ provided that the

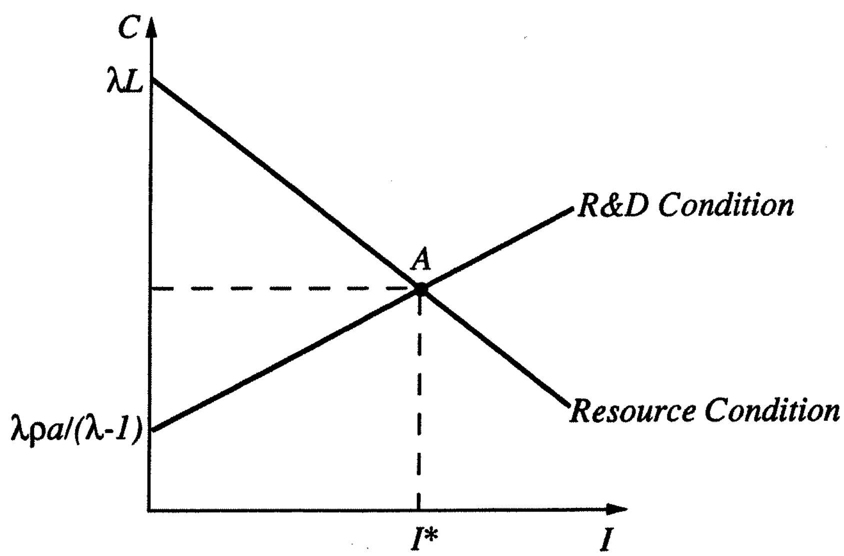

Figure 2: The Symmetric Steady State Equilibrium

economy's labor endowment is sufficiently large, that is, $L > \frac{a\rho}{\lambda-1}$. I will assume that this is the case, since otherwise growth does not occur. This completes the description of the model.

As illustrated in Figure 2, the R&D condition is globally upward sloping and the labor market condition is globally downward sloping. The upward slope of the R&D condition has a simple intuitive explanation: When consumer expenditure C increases, firms earn higher profit flows from winning R&D races [see (4)]. Without any increase in the industry level R&D intensity I, the marginal benefit from hiring one more worker to do R&D ($\frac{\pi L}{a(\rho+I)}$) would exceed the marginal cost (wage=1) for each firm. Thus profit maximizing firms respond by doing more R&D and I increases.[13] The downward sloping property of the labor market condition has an even simplier explanation: When consumer expenditure C increases, firms must employ more labor in production to satisfy the increased consumer demand and with a finite total endowment of labor L, there is less labor in the economy that can be devoted to R&D activities.

The steady state equilibrium given by point A in Figure 2 has a distinctively Schumpeterian flavor. Successful innovators replace previous industry leaders and snatch from them a 100% share of industry profits. The randomness of R&D success implies that progress occurs unevenly in each industry. Firms continually race to bring out the next generation products but there may be long periods without a success in some industries. However, all the uncertainty at the micro level washes out at the macro level and the economy grows at a smooth constant rate.

If the government subsidizes each R&D race at the rate $s_R > 0$ and finances the subsidy through lump-sum taxation, then the only change in the steady state conditions is that the RHS of (6) becomes $a(1-s_R)$. It follows that an increase in the R&D subsidy s_R causes the R&D condition in Figure 2 to shift down, resulting in a higher steady state R&D intensity I^* in each industry. Thus R&D subsidies serve to stimulate technological change and economic growth.[14] In this model, the growth rate is endogenously determined and public policies influence this growth rate.

[13]The upward sloping R&D condition is sometimes called the *Schumpeter curve* because it embodies the notion that innovation is driven by the quest for profit opportunities.

[14]Supporting this property of the model, Coe and Helpman [1993] demonstrate that investment in R&D is highly correlated with productivity growth in a sample of 22 OECD countries (a group that has relatively satisfactory data on R&D expenditures). Also Lach and Shankerman [1989] provide evidence that industrial research may be the primitive force behind much of the output growth that accounting methods attribute to factor accumulation. They find that R&D Granger-causes capital investment.

Another interesting implication of Figure 2 is called the "scale effect" property. Since an increase L causes the labor market condition to shift out, growth occurs at a more rapid pace when the economy has a larger resource base. Rivera-Batiz and Romer [1991] have emphasized the implications of this "scale effect" property for international trade. If one takes two structually identical countries that are completely isolated, then economic integration (going from autarky to free international trade) has the same effect as doubling the resource base in each country. The scale effect property implies that economic integration stimulates long-run economic growth.

2.5 Equilibrium and Optimal Growth

To simplify the welfare analysis, I will assume in this section that all individuals have not only identical preferences, but also identical labor endowments and assets. I assume that each individual inelastically supplies one unit of labor. It follows that L also represents the number of workers in the economy and aggregate consumer expenditure satisfies $C = L \cdot c$.

Since along any steady state equilibrium path, consumers only buy state-of-the-art quality products, (2) can be rewritten as

$$\log u(t) = \int_0^1 \log \lambda^j d\omega + \log\left(\frac{c}{\lambda}\right) \tag{8}$$

where $j = j(\omega, t)$ is the state-of-the-art quality level in industry ω at time t. The index j increases when firms are successful in innovating, and firms engage in innovative R&D in all industries thoughout time in any steady state equilibrium. For any industry ω, the probability of exactly m improvements in a time interval of length τ is $f(m,\tau) = [I\tau]^m e^{-I\tau}/m!$ Thus $f(m,\tau)$ represents the measure of products that are improved exactly m times in an interval of length τ. Using the properties of the Poisson distribution (see Hoel, Port and Stone [1971], page 84), it follows that the integral in (8) equal

$$\sum_{m=0}^{+\infty} f(m,\tau)[\log \lambda^m] = tI \log \lambda \tag{9}$$

The growth rate of consumer utility g is obtained by differentiating (9) with respect to t:

$$g \equiv \frac{d \log u(t)}{dt} = \frac{\dot{u}(t)}{u(t)} = I \log \lambda. \tag{10}$$

Since R&D subsidies increase the innovation rate I, it follows that from (10) that R&D subsidies stimulate long-run economic growth as well.

Substituting for static consumer utility $\log u(t)$ in (1) using (8) and (9), I obtain steady state discounted utility for the representative consumer

$$U = \frac{1}{\rho}\log\frac{c}{\lambda} + \frac{\log\lambda}{\rho^2}I. \tag{11}$$

Discounted consumer utility is an increasing function of both consumer expenditure c and the economy's growth rate $I\log\lambda$. Maximizing (11) subject to the resource constraint (7) where $C = L \cdot c$ yields a welfare maximizing R&D intensity I_w

$$I_w = \frac{L}{a} - \frac{\rho}{\log\lambda} \tag{12}$$

In contrast, from (6) and (7), the equilibrium R&D intensity I_e (with $s_R = 0$) satisfies

$$I_e = \frac{\lambda - 1}{a\lambda}L - \frac{\rho}{\lambda}. \tag{13}$$

Comparing (12) and (13), it is optimal to tax R&D if and only if

$$I_w - I_e = \frac{\rho}{\lambda}\left(\frac{L}{\rho a} + 1 - \frac{\lambda}{\log\lambda}\right) < 0 \tag{14}$$

Since $\lim_{\lambda\to+\infty}\frac{\lambda}{\log\lambda} = \lim_{\lambda\to 1}\frac{\lambda}{\log\lambda} = +\infty$, it is optimal to tax R&D if the size of innovations is either very large or very small. However, a more convenient way to interpret (14) is that, for any $\lambda > 1$, $I_w - I_e$ is positive if the labor force L is sufficiently large. When the labor force is relatively large and economic growth is fast, it is desirable to subsidize R&D so that economic growth becomes even faster. On the other hand, when the labor force is relatively small and economic growth is slow, it is desirable to tax R&D so economic growth becomes even slower. If the economy is growing slowly, it is growing too fast and if the economy is growing rapidly, it is not growing rapidly enough!

To understand the intution behind these seemingly paradoxical welfare properties, it is helpful to think about the externalities identified by Aghion and Howitt [1992]. Every time a firm innovates, consumers benefit because they can buy a higher quality product at the same price that they used to pay for a lower quality substitute. Furthermore these consumer benefits last forever because future innovations build on all the innovations of the past. Counter balancing this positive consumption externality is a negative business-stealing externality. Every time a firm innovates, it drives another firm out of business, destroying the other firm's profits. Individual R&D firms do not take into account these external losses in their profit maximization calculations whereas a social planner would.

When the economy's labor force is large, firms devote a lot of resources to R&D in equilibrium, the economy grows rapidly, and innovators typically earn quality leader profits for a short period of time (they are driven out of business quickly by further innovation). Because innovative firms are short-lived, the positive consumption externality dominates the negative business-stealing externality, and in the absence of government intervention, R&D effort is insufficient.

On the other hand, when the economy's labor force is small, it is not profitable for firms to do much R&D in equilibrium, the economy grows slowly, and innovators typically earn quality leader profits for a long period of time. Because innovative firms are long-lived, the negative business-stealing externality dominates the positive consumption externality, and in the absence of government intervention, R&D effort is excessive.

3 Segerstrom [1996]

One property of the model presented in the previous section has come in for a good deal of criticism recently: the larger is the economy's labor force (L), the higher is the steady-state growth rate (g). Since the world population growth rate has been positive through most of recorded history (see Kremer [1993]), this model implies that one should observe an upward trend in economic growth rates over time as the world economy becomes larger. Romer [1986] finds that there has been a steady upward trend in economic growth rates for technological leader countries from the 1700's until 1960. However, since 1960, per capita growth rates have exhibited either a constant mean or have declined on average (see Jones [1995a]). This is disturbing because during the same time period, the resources devoted to R&D increased substantially. For example, the number of scientist and engineers engaged in R&D in the United States increased from under 500,000 in 1965 to nearly 1 million in

1989 (see Table 4).[15] Given that recent experience is not consistent with "scale effect"

Table 4: Scientists And Engineers Engaged In R&D (thousands)

	U. S.	Japan	W. G.	France	U. K.
1965	494.2	117.6	61.0	42.8	49.9
1975	527.4	225.2	103.7	65.3	80.5
1985	841.2	381.3	143.6	102.3	97.8
1989	949.3	461.6	176.4	120.7	NA

property, Jones [1995b], Young [1995] and Segerstrom [1996] have developed R&D-driven growth models where scale effects are not present.

In this section, I present the knowledge creation model developed by Segerstrom [1996] to illustrate how the theory has evolved over time. The model is almost identical to the quality ladders model described in the previous section and thus, in describing the model, I will focus on the differences. There are two key differences. First, the population of workers is assumed to grow over time. Second, the returns to investing in R&D change over time. As time passes, innovating becomes progressively more difficult.[16] In Grossman and Helpman [1991], there is no population growth and the return to investing in R&D does not change over time in any industry. There is also one minor difference. Instead of deriving consumer behavior from individual utility maximization calculations, consumer behavior is derived from household (or family) utility maximization calculations. In both models, there is a

[15]National Science Board [1993], Appendix table 3-22.

[16]The primary motivation for this assumption change comes from the evidence on patenting behavior. As is noted by Kortum [1996], the number of patents per researcher has clearly declined in the United States during the last 40 years. Furthermore, this decline is not a recent development or limited to the United States. Machlup [1962] shows that the patents per researcher ratio declined consistently from 1920 to 1960 in the U.S., and Evenson [1984] shows that the decline in patents per researcher is a world-wide phenomenon. I interpret the decline in the number of patents per researcher as evidence that patentable inventions have become increasingly difficult to discover, as does Kortum [1996]. There are other possible interpretations (see Griliches [1990]).

continuum of industries, each industry has the same quality ladder structure, firms invest in R&D to discover higher quality products and R&D investment by firms is the engine driving economic growth.

3.1 Consumers and Workers

The economy has a fixed number of identical households that provide labor services in exchange for wages, and save by holding assets of firms engaged in R&D. Each individual member of a household is endowed with one unit of labor, which is inelastically supplied. The number of members in each family grows over time at the exogenous rate $n > 0$. Without loss of generality, I normalize the total number of individuals in the economy at time 0 to equal unity. Then the population of workers in the economy at time t is $L(t) = e^{nt}$. Each household is modelled as a dynastic family[17] which maximizes the discounted utility

$$U \equiv \int_0^\infty e^{nt} e^{-\rho t} \log u(t)\, dt \tag{15}$$

where $\rho > 0$ is the common subjective discount rate and $\log u(t)$ is the utility flow per person at time t, which is given by (2).

At each point in time t, each household allocates expenditure to maximize $\log u(t)$ given the prevailing market prices. As was the case earlier, solving this optimal control problem yields a unit elastic demand function ($d = c/p$ where d is quantity demanded, c is individual consumer expenditure and p is the relevant market price) for the product in each industry with the lowest quality adjusted price. The quantity demanded for all other products is zero. To break ties, I assume that when quality adjusted prices are the same for two products of different quality, a consumer only buys the higher quality product.

Given this static demand behavior, the intertemporal maximization problem of the representative household is equivalent to

$$\max_{c(t)} \int_0^\infty e^{-(\rho - n)t} \log c(t)\, dt \tag{16}$$

[17]Barro and Sala-i-Martin [1995, chapter 2] provide more details on this formulation of the household's behavior within the context of the Ramsey model of growth.

subject to the intertemporal budget constraint $\dot{a}(t) = w + r(t)a(t) - c(t) - na(t)$, where $a(t)$ denotes the per capita financial assets at time t, w is the wage income of the representative household member, and $r(t)$ is the instantaneous rate of return. Solving this intertemporal maximization problem yields the same differential equation as before, namely, the intertemporal optimization condition (3).

3.2 R&D Races

Labor is the only input used to do R&D in any industry, is perfectly mobile across industries and between production and R&D activities. There is free entry into each R&D race and all firms in an industry have the same R&D technology. Any R&D firm i that hires ℓ_i units of labor in industry ω at time t is successful in discovering the next higher quality product with instantaneous probability $\ell_i/X(\omega,t)$, where $X(\omega,t)$ is a R&D difficulty index. By instantaneous probability, I mean that $\frac{\ell_i}{X(\omega,t)}dt$ is the probability that the firm will innovate by time $t+dt$ conditional on not having innovated by time t, where dt is an infinitesimal increment of time. The returns to engaging in R&D races are independently distributed across firms, across industries, and over time. Thus, the industry-wide instantaneous probability of innovative success at time t is simply

$$I(\omega,t) \equiv \frac{L_I(\omega,t)}{X(\omega,t)} \tag{17}$$

where $\sum_i \ell_i = L_I$ is the industry-wide employment of labor in R&D.

The new and distinctive feature of this R&D technology is the $X(\omega,t)$ expression in the denominator of (17). I assume that R&D starts off being equally difficult in all industries $[X(\omega,0) = 1$ for all $\omega]$ and that R&D difficulty grows in each industry as firms do more R&D:

$$\frac{\dot{X}(\omega,t)}{X(\omega,t)} = \mu I(\omega,t), \tag{18}$$

where $\mu > 0$ is exogenously given. With this formulation, I capture in a simple way the idea that as the economy grows and $X(\omega,t)$ increases over time, innovating becomes more difficult. Ideas that are easier to discover tend to be discovered earlier, leaving more difficult ideas to be discovered later.[18] (17) also implies that a constant innovation rate I can be consistent

[18] I suppose that researchers can choose from a variety of different approaches to solving each problem

with positive growth in R&D labor employment L_I when X grows over time. Thus, this model has the potential to explain the evidence cited at the beginning of this section: growth in the number of scientists and engineers engaged in R&D without accelerating economic growth.

Let $v(\omega,t)$ denote the expected discounted profit or reward for winning a R&D race (in industry ω at time t) and s_R denote the fraction of each firm's R&D costs paid by the government. I will assume that the government finances the chosen R&D subsidy s_R using lump-sum taxation. Then at each point in time t, each R&D firm i chooses it's labor input ℓ_i to maximize its expected profits $\frac{v(\omega,t)\ell_i}{X(\omega,t)} - \ell_i(1-s_R)$. If $v(\omega,t) > X(\omega,t)(1-s_R)$, then $\ell_i = +\infty$ is profit maximizing and if $v(\omega,t) < X(\omega,t)(1-s_R)$, then $\ell_i = 0$ is profit maximizing. Only when

$$v(\omega,t) = X(\omega,t)(1-s_R) \tag{19}$$

is it profit maximizing for firms to devote a positive (finite) amount of labor to R&D. (19) implies that when R&D is more difficult [$X(\omega,t)$ is higher] or when the government subsidizes R&D less [s_R is lower], the reward for winning a R&D race must be larger to induce positive R&D effort. When (19) holds, firms are globally indifferent concerning their choice of R&D effort. Given the symmetric structure of the model, I focus on equilibrium behavior where the R&D intensity $I(\omega,t)$ is the same in all industries ω at time t and is strictly positive. Thus the ω argument of functions is dropped in the rest of this section.

The stock market valuation of monopoly profits provides another equilibrium condition that relates the expected discounted profits to the flow of profits and the instantaneous interest rate. Over a time interval dt, the shareholder receives a dividend $\pi^L(t)\,dt$, and the value of the monopolist appreciates by $\dot{v}(t)\,dt$ in each industry. Because each quality leader is targeted by other firms that conduct R&D to discover the next higher quality product, the shareholder suffers a loss of $v(t)$ if further innovation occurs. This event occurs with probability $I(t)\,dt$, whereas no innovation occurs with probabillity $[1-I(t)]\,dt$. Efficiency in

that they face. Those approaches with the highest ex ante probability of success are pursued first and when success does not materialize, researchers switch to pursuing ex ante less promising approaches to solving problems. Thus, those innovations that are easier to discover tend to be discovered earlier in time, leaving the more difficult innovations to be discovered later.

financial markets requires that the expected rate of return from holding a stock of a quality leader is equal to the riskless rate of return $r(t)\,dt$ that can be obtained through complete diversification: $\frac{\pi^L}{v}dt + \frac{\dot{v}}{v}(1 - I\,dt)\,dt - \left[\frac{v-0}{v}\right] I\,dt = r\,dt$. Taking limits as dt approaches zero yields:

$$v(t) = \frac{\pi^L(t)}{r(t) + I(t) - \frac{\dot{v}(t)}{v(t)}}. \tag{20}$$

The profits earned by each leader π^L [given by (4) where $C(t) = c(t)L(t)$] are appropriately discounted using the interest rate r and the instantaneous probability I of being driven out of business by further innovation. Also taken into account in (20) is the possibility that these discounted profits grow over time. R&D profit maximization implies that the discounted marginal revenue product of an idea (20) must equal its marginal cost (19) at each point in time, that is,

$$x(t)(1 - s_R) = \frac{\left(\frac{\lambda-1}{\lambda}\right) c(t)}{r(t) + (1-\mu)I(t)}, \tag{21}$$

where $x(t) \equiv \frac{X(t)}{L(t)}$. Note that (18) and (19) imply that $\frac{\dot{v}(t)}{v(t)} = \frac{\dot{X}(t)}{X(t)} = \mu I(t)$.

3.3 The Labor Market

In each industry ω at time t, consumers only buy from the current quality leader and pay the equilibrium price λ. Since consumer demand is unit elastic, $c(t)L(t)/\lambda$ workers must be employed by the current quality leader to produce enough to meet consumer demand. In addition, $L_I(t)$ workers are employed by R&D firms in each industry at time t. With a measure one of industries, full employment of workers is satisfied when $L(t) = \frac{c(t)L(t)}{\lambda} + L_I(t)$ holds for all t. Substituting using $x(t) \equiv \frac{X(t)}{L(t)}$ and $L_I(t) = I(t)X(t)$ into the full employment condition yields

$$1 = \frac{c(t)}{\lambda} + I(t)x(t). \tag{22}$$

At any point in time t, more consumer expenditure $c(t)$ comes at the expense of less R&D investment $I(t)$.

3.4 Balanced Growth Equilibria

I now solve the model for balanced growth equilibrium paths where all endogenous variables grow at constant (not necessarily the same) rates and firms invest in R&D [$I(t) > 0$ for all t].

Given $\mu > 0$, (18) implies that I must be a constant over time. It then follows from (22) that c and x must also be constants. Thus, any balanced growth equilibrium must involve c, x and I taking on constant values over time.

Differentiating (17) with respect to time using (18) yields $\frac{\dot{I}}{I} = \frac{\dot{L}_I}{L_I} - \mu I = 0$. (17) and (22) together imply that in any balanced growth equilibrium, employment in the R&D sector $L_I(t)$ must grow at the same rate as the population (n). Thus, there is a unique balanced growth R&D intensity

$$I = \frac{n}{\mu}. \tag{23}$$

The level of R&D investment is completely determined by the exogenous rate of population growth $n > 0$ and the R&D difficulty growth parameter $\mu > 0$. The balanced growth innovation rate is higher when the population of consumers grows more rapidly or when R&D difficulty increases more slowly over time. Note that if there is no increase in R&D difficulty over time ($\mu \leq 0$), then there is no balanced growth equilibrium. Instead, the growth rate of the economy increases without bound over time.

Many economists react negatively to the implication of (23) that population growth is good for economic growth. They point to the many empirical studies which find that countries with faster population growth rates tend to experience lower per capita income growth. However, this reaction shows a misunderstanding of the theory. In the model, the relevant population growth rate is the growth rate of consumers that a quality leader sells to. In a world where countries are closely linked by international trade, it makes more sense to interpret the population growth rate in the model n as the *world* population growth rate than the population growth rate of any particular country. Thus, a more appropriate test of the theory is that world economic growth is higher during periods of time when the world population growth rate is higher. According to Kremer [1993], the world population growth rate has been increasing throughout most of recorded history (see Table 5) and according

Table 5: World Population Growth: 1700-1990

Period	Growth Rate	Period	Growth Rate
1700-1750	0.33%	1930-1940	1.07%
1750-1800	0.44%	1940-1950	1.28%
1800-1850	0.57%	1950-1960	1.82%
1850-1875	0.39%	1960-1970	2.01%
1875-1900	0.81%	1970-1980	1.86%
1900-1920	0.83%	1980-1990	1.81%
1920-1930	0.91%		

to Romer [1986], there has been a significant upward trend since 1700 in the rate of growth of output per person-hour in the world's highest productivity countries[19] (see Table 6). Thus, this particular implication of the theory should not be so quickly dismissed as being

Table 6: Productivity Growth Rates for Leading Countries

Lead Country	Interval	Growth Rate
Netherlands	1700-1785	-0.07%
United Kingdom	1785-1820	0.5%
United Kingdom	1820-1890	1.4%
United States	1890-1979	2.3%

inconsistent with the empirical evidence.[20]

[19]As Romer explains, "Growth for a country that is not a leader will reflect at least in part the process of imitation and transmission of existing knowledge, whereas the growth rate of the leader gives some indication of growth at the frontier of knowledge." In Table 6, the growth rates reported are annual average compound growth rates of GDP per man-hour.

[20]It is interesting that the recent period of declining world population growth (since 1970) is also the period of time where researchers have identified a productivity slowdown. This may be a pure coincidence but needs to be investigated further. The theory predicts that, other things being equal, a permanent decline

With I given by (23), and (3) implying that the equilibrium interest rate is $r(t) = \rho$, (21) yields a balanced growth R&D condition

$$(1 - s_R)x = \frac{\left(\frac{\lambda-1}{\lambda}\right) c}{\rho + \frac{n}{\mu} - n}, \tag{24}$$

and (22) yields a balanced growth resource condition

$$1 = \frac{c}{\lambda} + \frac{nx}{\mu}. \tag{25}$$

Both balanced growth conditions are illustrated in Figure 3. The vertical axis measures

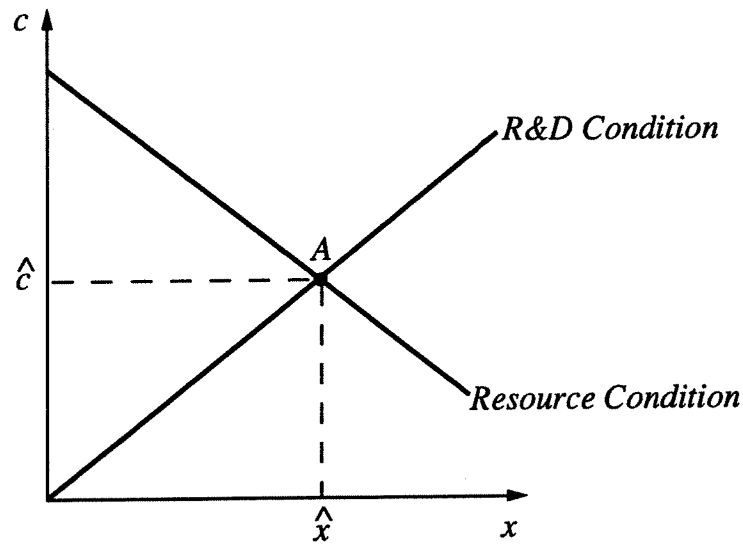

Figure 3: The unique balanced growth equilibrium

consumption per capita c and the horizontal axis measures relative R&D difficulty x. The R&D condition is upward sloping in (x, c) space, indicating that when R&D is relatively more difficult, consumer expenditure must be higher to justify positive R&D effort by firms. The resource condition is downward sloping in (x, c) space, indicating that when R&D is relatively more difficult and more resources are used in the R&D sector to maintain the balanced growth innovation rate I, less resources are available to produce goods for consumers, so individual consumers must buy less. The unique intersection between the R&D and resource conditions at point A determines the balanced growth values of consumption per capita $\hat{c}$ and relative R&D difficulty $\hat{x}$.

is the rate of population growth would lead to a permanent slowdown in the rate of economic growth.

In the balanced growth equilibrium, the value of patented innovations rises over time ($\frac{\dot{v}(t)}{v(t)} = \mu I = n$) causing firms to expend ever greater resources to discover them ($\frac{\dot{L}_I(t)}{L_I(t)} = n$). Although the patents per researcher ratio ($\frac{1}{X(t)}$) falls over time, the value of patents discovered per researcher ($\frac{v(t)}{X(t)}$) remains constant over time.

If $x = \hat{x}$ at time $t = 0$, then an immediate jump to the balanced growth path can occur. Otherwise, it is imperative to investigate the transitional dynamic properties of the model.

Differentiating relative R&D difficulty $x(t) \equiv \frac{X(t)}{L(t)}$ with respect to time using (18) yields $\frac{\dot{x}(t)}{x(t)} = \mu I(t) - n$. Substituting into this expression for $I(t)$ using the the resource condition (22) yields one differential equation that must be satisfied along any equilibrium path for the economy:

$$\dot{x}(t) = \mu\left(1 - \frac{c(t)}{\lambda}\right) - nx(t) \qquad (26)$$

Since the RHS of (26) is decreasing in both x and c, $\dot{x}(t) = 0$ defines the downward-sloping curve in Figure 4. Starting from any point on this curve, an increase in x leads to $\dot{x} < 0$ and

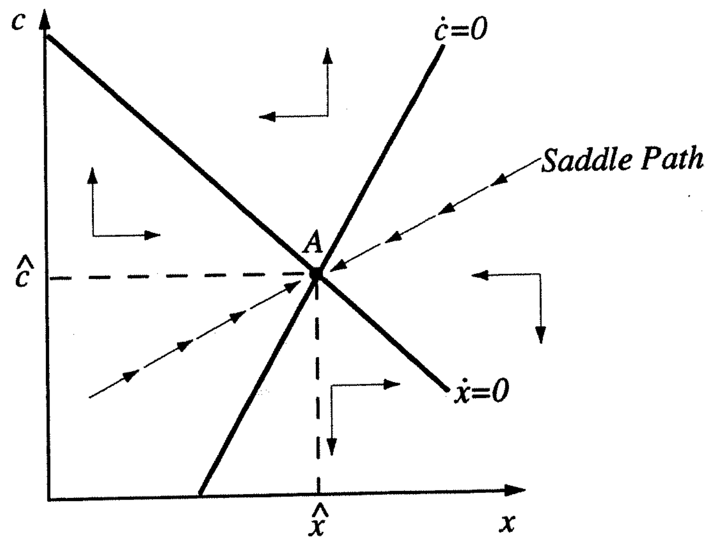

Figure 4: Stability of the balanced growth equilibrium

a decrease in x leads to $\dot{x} > 0$, as is illustrated by the horizontal arrows in Figure 4. Solving (22) for $I(t)$, (21) for $r(t)$ and then substituting into (3) yields a second differential equation that must be satisfied along any equilibrium path for the economy:

$$\dot{c}(t) = c(t)\left[\frac{(\lambda - 1)}{\lambda(1 - s_R)}\frac{c(t)}{x(t)} + \frac{(\mu - 1)}{x(t)}\left(1 - \frac{c(t)}{\lambda}\right) - \rho\right] \qquad (27)$$

If $\mu \leq 1$, then the $\dot{c}(t) = 0$ curve is definitely upward sloping in (x, c) space. Starting from any point on this curve, an increase in x leads to $\dot{c} < 0$ and a decrease in x leads to $\dot{c} > 0$, implying that there exists an upward-sloping saddle path. If μ is slightly greater than 1, then the $\dot{c}(t) = 0$ curve is still upward sloping in (x, c) space and there exists an upward sloping saddle path (this case is illustrated in Figure 4). Even if μ is significantly greater than 1 and the $\dot{c}(t) = 0$ curve is downward sloping, there still exists an upward-sloping saddle path going through the unique balanced growth equilibrium point A. Thus the balanced growth equilibrium is saddle path stable. By jumping onto this saddle path and staying on it forever, convergence to the balanced growth equilibrium occurs, just like in the neoclassical growth model.

Along a balanced growth path, the fraction of the labor force devoted to R&D is uniquely determined. (17) implies that $\frac{L_I(t)}{L(t)} = Ix$. Solving (24) and (25) for x yields a final balanced growth condition:

$$\frac{L_I}{L} = \frac{1}{1 + \frac{(1-s_R)}{(\lambda-1)}\left[1 + \frac{\rho - n}{I}\right]} \tag{28}$$

Given (23), the balanced growth fraction of the labor force devoted to R&D $\frac{L_I}{L}$ is completely determined by parameter values. Interestingly, although a higher R&D subsidy has no effect on the long-run innovation rate, it does increase the fraction of workers in the economy doing R&D.

The share of labor devoted to R&D has been changing over time in several advanced countries.[21] Throughout the relevant time period (1965-1989), the United States had the highest fraction of workers engaged in R&D. In all the other countries, the share of labor devoted to R&D has steadily increased over time, with the biggest increase occurring in Japan. Part of the increase in these countries can be attributed to convergence to American levels but even the share of labor devoted to R&D in the United States has increased, particularly after 1975.

In Dinopoulos and Segerstrom [1996], one possible explanation for this trend is provided, namely, trade liberalization. The international effort to cut tariff and nontariff barriers embodied in GATT, NAFTA, WTO and other agreements has contributed to an explosion

[21]Source: National Science Board [1993].

in international trade. In the United States, imports and exports were about 3 percent of GDP in 1970, as opposed to 10-12 percent today. According to World Bank figures, between 1965 and 1990, the share of output exported rose for low-income countries from 8 to 18 percent, for middle-income countries from 17 to 25 percent and for high-income countries, from 12 to 20 percent.[22] Using a two country version of the model described in this section, Dinopoulos and Segerstrom show that lower tariff barriers cause the share of labor devoted to R&D to increase in both trading countries.

3.5 Utility Growth

Along a balanced growth path, the representative consumer's utility grows at a constant rate. To solve for this utility growth rate, I substitute the static consumer demand for quality leader products $d(j,\omega,t) = c/\lambda$ into (2) to obtain

$$\log u(t) = \log c(t) - \log \lambda + \log Q(t) \tag{29}$$

where $Q(t)$ is defined by $\log Q(t) \equiv \int_0^1 \log \lambda^{j(\omega,t)} d\omega$. Since the measure of industries with exactly m innovations at time t is $(It)^m e^{-It}/m!$, $\log Q(t) = \sum_{m=0}^{\infty} \frac{(It)^m e^{-It}}{m!} \log \lambda^m = It \log \lambda$. Thus, differentiation of (29) yields

$$\frac{\dot{u}}{u} = I \log \lambda. \tag{30}$$

Individual utility growth depends on the rate at which new higher quality products are introduced. Treating individual utility growth as the measure of economic growth for the economy, (30) implies that a higher R&D subsidy has no effect on economic growth since it does not impact the innovation rate in any industry.

3.6 Optimal Growth

In this subsection, I explore the balanced growth properties of the model when all allocation decisions are made by a social planner. I assume that the social planner's objective is to maximize the discounted utility of the representative family.

[22]See Richardson [1995].

I will first show that the social planner chooses the same R&D intensity in each industry. Allowing for different R&D intensities in different industries, the labor market constraint is $L(t) = \frac{c(t)L(t)}{\lambda} + \int_0^1 L_I(\omega, t)\, dt$. Thus the quantity consumed of each state-of-the-art quality product by the representative consumer at time t is $d(t) \equiv 1 - \int_0^1 \frac{L_I(\omega,t)}{L(t)} d\omega$. For given resources devoted to R&D at time t, $\int_0^1 L_I(\omega, t) d\omega$, clearly the social planner wants to choose the distribution of R&D expenditures across industries to maximize $\frac{\dot{Q}(t)}{Q(t)} = \int_0^1 I(\omega, t) \log \lambda\, d\omega = \int_0^1 \frac{L_I(\omega,t)}{X(\omega,t)} \log \lambda\, d\omega$ [where I have used (17)]. This implies that at any point in time t, all R&D is done in those industries with the lowest $X(\omega, t)$ and with the social planner carrying out such a policy throughout time, $X(\omega, t) = X(t)$ and $I(\omega, t) = I(t)$ for all ω and t.

Substituting the appropriately simplified expression $d(t) = 1 - \frac{X(t)I(t)}{L(t)}$ back into the discounted utility function (15), I can solve for the discounted utility of the representative family. The optimal control problem facing the social planner is given by

$$\max_{I(\cdot)} \int_\tau^\infty e^{(n-\rho)t} \left\{ \log Q(t) + \log \left[1 - \frac{I(t)X(t)}{L(t)} \right] \right\} dt \tag{31}$$

subject to two state equations

$$\dot{X}(t) = X(t)I(t)\mu, \tag{32}$$

$$\dot{Q}(t) = Q(t)I(t) \log \lambda, \tag{33}$$

two initial conditions $[X(\tau), Q(\tau) > 0$ given$]$, one control constraint $[I(t) \geq 0$ for all $t]$, and the population growth condition $[N(t) = e^{nt}]$.

This optimal control problem is solved in the Appendix of Segerstrom [1996]. I find that there is a unique balanced growth solution:

$$I = \frac{n}{\mu} \tag{34}$$

and

$$\frac{L_I}{L} = \frac{1}{1 + \frac{\mu}{\log \lambda} + \frac{\rho - n}{I \log \lambda}}. \tag{35}$$

Even though the optimal innovation rate (34) coincides with the equilibrium innovation rate (23), the government can in general improve welfare by intervening in the economy. Comparing (35) with (28), the optimal R&D subsidy rate s_R must satisfy

$$(1 - s_R)\frac{\log \lambda}{\rho - n} = \frac{\lambda - 1}{I + \rho - n} \left[1 + \frac{I\mu}{\rho - n} \right]. \tag{36}$$

Since $\frac{\lambda-1}{\log \lambda}$ is a globally increasing function of λ and the equilibrium innovation rate I given by (23) does not depend on λ, (36) implies that the optimal R&D subsidy rate s_R is a globally decreasing function of λ, holding all other parameter values fixed.[23] Since $\lim_{\lambda\to\infty} \frac{\log \lambda}{\lambda-1} = 0$, (36) also implies that it is optimal to tax R&D if λ is sufficiently large. Finally, holding all other parameters fixed, (36) implies that it is optimal to subsidize R&D if $\mu > 0$ is sufficiently small. Thus, both R&D subsidies and R&D taxes can be optimal, depending on the parameters of the model.

To understand the intuition behind these welfare results, it is helpful to think about the following three externalities: Every time a firm innovates, consumers benefit because they can buy a higher quality product at the same price that they used to pay for a lower quality substitute. Furthermore these consumer benefits last forever because future innovations build on all the innovations of the past. This positive *consumption effect* measures $\frac{\log \lambda}{\rho-n}$ in terms of the utility metric given by (15) and (2). Counter balancing this positive consumption effect is a negative *business-stealing effect.* Every time a firm innovates, it drives another firm out of business, destroying the other firm's profits. Individual R&D firms do not take into account these external losses in their profit maximization calculations but they are taken into account by a social planner. The size of this business stealing externality is given by the first expression on the right hand side of (36). The term $\lambda - 1$ represents the per unit profit flow earned by a quality leader. This profit flow is discounted by $I + \rho - n$ since each firm is eventually driven out of business by further innovation and profits are expected to grow until then. Finally, there is a negative *intertemporal spillover effect.* By innovating today, a firm raises the costs of all R&D firms in the future. In deciding how much R&D to do, a firm only focuses on present costs and benefits but a social planner takes into account how R&D costs increase as a result of present investment decisions. Since R&D success today implies that future business creation becomes permanently more difficult, this negative spillover effect is given by the $\frac{\mu I}{\rho-n}$ term in (36). In the limiting case where $\mu = 0$, R&D does not become more difficult over time and the intertemporal spillover effect disappears.

[23]The proof of the first claim is as follows: First I note that $\lambda \cdot \log \lambda > \lambda - 1$ for all $\lambda > 1$ since $(\lambda \cdot \log \lambda)' = 1 + \log \lambda > (\lambda - 1)' = 1$ and equality holds when $\lambda = 1$. Therefore, using the fact that $\log \lambda > \frac{\lambda-1}{\lambda}$ holds for all $\lambda > 1$, differentiation reveals that $\frac{\lambda-1}{\log \lambda}$ is a increasing function of λ for all $\lambda > 1$.

When λ is large, that is, new products represent big improvements over existing products, then innovative firms are able to charge big markups of price over marginal cost and earn big profits. Under these circumstances, the negative business-stealing effect associated with R&D dominates. Innovative firms do not take into account in their profit-maximizing calculations the large losses in discounted profits incurred by the firms they drive out of business. In the pursuit of other firms' profits, too large a fraction of the economy's resources are devoted to R&D and a R&D tax is welfare-maximizing. On the other hand, when μ is small, that is, innovation only becomes slightly more difficult over time, then balanced growth for the economy is associated with innovations occurring frequently. Then innovative firms earn quality leader profits for a short period of time before they are driven out of business by further innovation. Under these circumstances, the positive consumption effect associated with R&D dominates. Innovative firms only briefly benefit from their discoveries but the benefits to consumers last forever. Too small a fraction of the economy's resources are devoted to R&D and a R&D subsidy is welfare-maximizing.

In Grossman and Helpman [1991], there is a n-shaped relationship between the innovation size parameter λ and the optimal R&D subsidy s_R: R&D taxes are optimal for very small or very large size innovations but R&D subsidies are optimal for intermediate-size innovations. The main reason for this more complicated relationship is that the equilibrium innovation rate is an increasing function of λ in their model. Since I appears in the denominator of the business-stealing effect expression in (36), larger innovations lead to more firm turnover and this reduces the size of the business-stealing effect. In Segerstrom [1996], the equilibrium innovation rate only depends on the growth rate of the effective labor force. An increase in the size of innovations increases the relative size of the R&D sector but not the long-run growth rate of the effective labor force. Given that Segerstrom [1996] and Grossman and Helpman [1991] share many assumptions, I conclude that both the positive and normative properties of "quality ladders" growth models significantly change when less optimistic assumptions are made about the returns to R&D investment.

4 Mankiw, Romer and Weil [1992]

In this section, I present a model of physical and human capital accumulation developed by Mankiw, D. Romer and Weil [1992] to explain cross-country differences is incomes and growth rates. One of the nice features of the MRW model is that a special case of this model is the Solow [1956] growth model. Thus, I can compare how well models with human capital accumulation fit the data compared with "old growth theory" models that only have physical capital accumulation.

In MRW model, the production function for the economy at time t is given by

$$Y(t) = K(t)^{\alpha} \left(\frac{H(t)}{A(t)L(t)} \right)^{\beta} (A(t)L(t))^{1-\alpha} \tag{37}$$

where $Y(t)$ is output, K(t) is physical capital, $H(t)$ is human capital, $L(t)$ is labor, $A(t)$ represents "knowledge" or the "effectiveness of labor", and the production parameters satisfy $\alpha, \beta > 0$. $A(t)L(t)$ is the *effective labor* used in production and $\frac{H(t)}{A(t)L(t)}$ represents the *human capital per effective unit of labor.* A skilled worker supplies both 1 unit of L and some amount of H. (37) states that output depends on the quantities of physical capital (K) and effective labor (AL) used in production, as well as on the "average" level of human capital posessed by the labor force (H/AL). Furthermore, holding fixed the level of human capital per effective worker, there is constant returns to scale (output doubles when both K and AL are doubled). MRW assume that $\alpha + \beta < 1$, so there are diminishing returns to capital accumulation broadly defined.

L and A are assumed to grow at exogenously given rates n and g:

$$L(t) = L(0)e^{nt} \tag{38}$$

$$A(t) = A(0)e^{gt}. \tag{39}$$

Since the goal of this model is not to explain worldwide growth but rather why there are income differences across countries, the rate of technological change g is taken as given. Constant fractions of income s_k and s_h are invested in physical capital and human capital, respectively. Both physical and human capital depreciate at the same rate δ. Thus, the dynamics of K and H are given by

$$\dot{K}(t) = s_k Y(t) - \delta K(t) \tag{40}$$

$$\dot{H}(t) = s_h Y(t) - \delta H(t). \tag{41}$$

MRW interpret (40) and (41) to mean that the technology for producing new physical (human) capital combines physical capital, human capital and raw labor in the same way as the technology for producing goods.

For analyzing the steady state properties of the model, it is convenience to define new endogenous variables $y \equiv \frac{Y}{AL}$, $k \equiv \frac{K}{AL}$, and $h \equiv \frac{H}{AL}$ as the levels of output, physical capital and human capital per effective unit of labor. Then the production function (37) can be rewritten more simply as

$$y(t) = k(t)^{\alpha} h(t)^{\beta} \tag{42}$$

and the differential equations describing physical and human capital accumulation become

$$\dot{k}(t) = s_k k(t)^{\alpha} h(t)^{\beta} - (\delta + n + g)k(t) \tag{43}$$

$$\dot{h}(t) = s_h k(t)^{\alpha} h(t)^{\beta} - (\delta + n + g)h(t) \tag{44}$$

In Figure 5, the phase diagram constructed from these two differential equations is illustrated. The model has a unique stable steady-state equilibrium given by point E. Regardless of where the economy starts in Figure 5, it converges to point E. Solving for when both $\dot{k}(t) = 0$ and $\dot{h}(t) = 0$, the steady-state levels of k and h can be determined:

$$k^* = \left(\frac{s_k^{1-\beta} s_h^{\beta}}{n + g + \delta} \right)^{1/(1-\alpha-\beta)} \tag{45}$$

$$h^* = \left(\frac{s_k^{\alpha} s_h^{1-\alpha}}{n + g + \delta} \right)^{1/(1-\alpha-\beta)} \tag{46}$$

There are three important things to note about this steady-state equilibrium. First, since k, h and y are constant over time, K, H and Y must grow at the same rate $n + g$ as the effective labor force AL. Thus, the growth rate of Y/L is g. The long-run growth rate of output per worker is determined by the exogenous rate of technological change. Parameters of the model like the population growth rate n, the physical capital savings rate s_k and the human capital savings rate s_h have no influence on the long run growth rate of the economy.

Second, both k^* and h^* are nontrivial functions of the savings and population growth rates (s_k, s_h and n). For example, the model predicts that, other things being equal, an

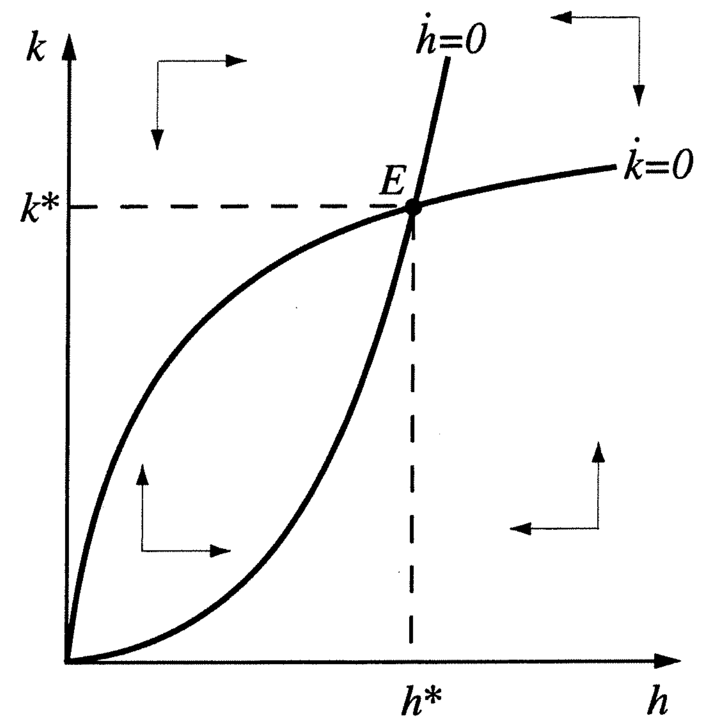

Figure 5: The dynamics of k and h

economy with a higher population growth rate will have lower steady state stocks of both physical and human capital per effective unit of labor. For another example, the model predicts that an economy with a higher human capital savings rate will have higher steady state stocks of both physical and human capital per effective unit of labor.

Third, the model predicts that two countries with the same savings and population growth rates will have the same steady state standards of living (at each point in time), even if one of the countries starts out at time $t = 0$ with twice as much physical and human capital per person. This property of the model is called "conditional convergence" in the literature. It can be eliminated by modifying the model (instead assuming that $\alpha + \beta = 1$ and $g = 0$). This modification yields a simplified version of the Lucas [1988] model where human capital accumulation is the engine of growth. It is straightforward to show that in this modified model, there exists a unique stable steady state equilibrium ratio h/k. Furthermore, conditional convergence does not definitely occur. For two countries with the same parameter values but different initial stocks of capital per worker, the initial differences can persist forever.

To understand the implications of the MRW model for steady-state income per capita, I substitute (45) and (46) into (42) and take logs to obtain

$$\log\left(\frac{Y(t)}{L(t)}\right) = \log A(0) + gt + \frac{\alpha\log(s_k) + \beta\log(s_h) - (\alpha+\beta)\log(n+g+\delta)}{1-\alpha-\beta} \tag{47}$$

Thus, in a steady-state equilibrium at time t, income per worker is an increasing function of s_k and s_h, and a decreasing function of n.

Table 7: Data For Selected Countries, 1960-1985

country	GDP/adult growth rate	1985 GDP per adult	pop. growth	invest. rate	school. rate
Singapore	6.6%	$14,678	2.6%	32.2%	9.0%
Japan	5.6	13,893	1.2	36.0	10.9
South Korea	5.2	4,775	2.7	22.3	10.2
Norway	3.6	19,723	0.7	29.1	10.0
Finland	2.9	13,779	0.7	36.9	11.5
West Germany	2.8	15,297	0.5	28.5	8.4
Sweden	2.7	15,237	0.4	24.5	7.9
United Kingdom	2.2	13,331	0.3	18.4	8.9
Mexico	2.2	7,380	3.3	19.5	6.6
United States	1.7	18,988	1.5	21.1	11.9
Philippines	1.5	2,430	3.0	14.9	10.6
Nigeria	1.4	1,186	2.4	12.0	2.3
India	1.2	1,339	2.4	16.8	5.1
Argentina	0.6	5,533	1.5	25.3	5.0
Zambia	-0.6	1,217	2.7	31.7	2.4
Madagascar	-0.8	975	2.2	7.1	2.6
Venezuela	-1.9	6,336	3.8	11.4	7.0

To the extent that savings and population growth rates differ across countries, the model predicts that there will be long-run differences in per capita incomes across countries. MRW test this implication of the model using data from the Real National Accounts constructed by Summers and Heston [1988]. The data are annual and cover the period 1960-1985. For selected countries, this data is displayed in Table 7. For each country, they measure n as the average rate of growth of the working-age population (ages 15 to 64), $Y(t)/L(t)$ as real GDP in 1985 divided by the working-age population in that year, and s_k as the average share of real investment (including government investment) in real GDP. As a proxy for the rate of human capital accumulation (s_h), they use the percentage of the working-age population that is in secondary school.[24]

For all countries, MRW assume that $g = 0.02$ (2% growth in income per capita) and $\delta = 0.03$ (3% depreciation rate). Rewriting (47) slightly, the equation that MRW estimate is

$$\log\left(\frac{Y_i}{L_i}\right) = a + b\log(s_{ki}) + c\log(s_{hi}) + d\log(n_i + 0.05) + \epsilon_i \tag{48}$$

where i indexes countries and ϵ_i is the country-specific error term.

If β is set equal to zero in (37) [which implies that $c = 0$ in (48)], then the MRW model becomes the Solow [1956] model. MRW first test how well the Solow model fits the data. Results are reported in the second and third columns of Table 8. Standard errors for all coeffecient estimates are in parentheses. For the broadest sample of 98 countries,[25] three aspects of the results support the Solow model. First, the coefficients on savings and population growth have the predicted signs and are highly significant. Second, the restriction imposed by the theory that $b+d = 0$ holds approximately and can not be rejected statistically. Third, the adjusted R^2 of 0.59 indicates that almost 60% of the cross-country variation in labor productivity can be accounted for by differences is population growth

[24]This proxy for the rate of human capital accumulation is crude not just because it ignores learning that occurs after secondary school (at universities, for example) but also because it measures quantity rather than quality. A recent study of the skills of average eigth graders in 41 countries found that Singapore and Japan were among the top 3 countries in both math and science tests. Interestingly, these countries are also among the fastest growing countries since 1960 (see Table 7). Unfortunately, measures of the quality of education have not yet been used by economists in cross country growth regressions.

[25]Not included in this list are countries for which oil production is the dominant industry and centrally planned economies.

Table 8: Estimation of the Steady-State Equation

Dependent variable: log GDP per working-age person in 1985				
Sample:	Non-oil	OECD	Non-oil	OECD
Observations:	98	22	98	22
Constant	5.48	7.97	6.89	8.63
	(1.59)	(2.48)	(1.17)	(2.19)
$\log(I/GDP)$	1.42	0.50	0.69	0.28
	(0.14)	(0.43)	(0.13)	(0.39)
$\log(n+g+\delta)$	-1.97	-0.76	-1.73	-1.07
	(0.56)	(0.84)	(0.41)	(0.75)
log(School)			0.66	0.76
			(0.07)	(0.29)
$\overline{R}^2$	0.59	0.01	0.78	0.24
Implied α	0.60	0.36	0.31	0.14
	(0.02)	(0.15)	(0.04)	(0.15)

rates and physical capital accumulation rates. Nevertheless, the "old growth theory" is not completely successful. The value of α implied by the coefficients should equal capital's share of income, which is roughly one third. Instead, it is 0.60 which is much too high and the standard error corresponding to this estimate is only 0.02. Furthermore, when attention is restricted to the smaller sample of 22 OECD countries (most West European countries plus the United States, Canada, Japan, Australia and New Zealand), the adjusted R^2 of 0.01 indicates that only 1% of the cross-country variation in labor productivity within OECD countries can be accounted for by differences is population growth and investment rates! MRW conclude that "all is not right with the Solow model."

MRW next test how well their own model with human capital accumulation fits the data. Their results are reported in columns 3 and 4 of Table 8. For the broadest sample 98 countries, the results are surprisingly good. First, the coefficients on population growth and both types of savings (physical and human capital accumulation) have the predicted signs and are highly significant. Second, the restriction imposed by the theory that $b+c+d=0$ holds approximately and can not be rejected statistically. Third, the adjusted R^2 of 0.78 indicates that almost 80% of the cross-country variation in labor productivity can be accounted for by differences is population growth rates as well as physical and human capital accumulation rates. Fourth, the value of α implied by the coefficients roughly equals one third, as it should. Considering that an imprecise proxy for human capital is used (the proportion of the working age population that is in secondary school), these results are remarkably good, and suggest that the Solow model only needs to be augmented by allowing for human capital accumulation.

However, when attention is restricted to the smaller sample of 22 OECD countries, a very different picture emerges. First, although the coefficients on population growth and both types of savings (physical and human capital accumulation) have the predicted signs, only the human capital coefficient is statistically significant. Second, the adjusted R^2 of 0.24 indicates that less than 25% of the cross-country variation in labor productivity can be accounted for by differences is population growth and savings rates. It appears that the high adjusted R^2 obtained for the large sample of 98 countries is mostly driven by differences in investment ratios and population growth rates between rich and poor countries. The MRW model does not appear to do a good job of explaining income differences between OECD countries.

Next, MRW turn to the issue of how well their model with physical and human capital accumulation can account for differences in *growth rates* across countries. In the process, they provide one interpretation for why their model does a poor job of accounting for income differences between OECD countries.

In a steady state equilibrium, per capita income in all countries grows at the same rate g. However, outside of the steady-state, countries can experience different growth rates. The convergence illustrated in Figure 5 just means that these growth rate differences diminish with time. Taking logs of both sides of (42) and then differentiating with respect to time yields

$$\frac{d}{dt}[\log y(t) - \log y^*] = \alpha\frac{\dot{k}(t)}{k(t)} + \beta\frac{\dot{h}(t)}{h(t)} \approx -\lambda[\log y(t) - \log y^*] \tag{49}$$

where $\lambda \equiv (n+g+\delta)(1-\alpha-\beta) > 0$. The second line in (42) is obtained by taking first-order Taylor approximations of $\dot{k}$, $\dot{h}$, $\log k$ and $\log h$ around the steady-state equilibrium using (43), (44), (45) and (46). Solving the simple differential equation given by the second line in (50) yields

$$\log y(t) \approx e^{-\lambda t} \log y(0) + (1 - e^{-\lambda t}) \log y^* \tag{50}$$

Equation (50) has a strong implication, namely, that convergence to the steady state is much slower in the MRW model (with human capital accumulation) than in the Solow model (without human capital accumulation). Thus the MRW model has more potential for explaining differences in growth rates across countries. To illustrate, suppose that $\alpha = \beta = 1/3$ (these are close to the estimated values for the 98 country sample), $n = 0.01$ (1% population growth), and $g+\delta = 0.05$. For these reasonable parameter values, the convergence rate is $\lambda = 0.02$ and (50) implies that the economy moves halfway to the steady-state in roughly 35 years.[26] In constrast, if β is set equal to zero (the Solow model) while leaving all other parameter choices unchanged, then convergence is twice as fast ($\lambda = 0.04$) and the economy moves halfway to the steady-state in roughly 17 years.

Substituting for y^* in (50) using (42), (45) and (46) yields

$$\begin{aligned} \log\left[\frac{Y(t)}{L(t)}\right] - \log\left[\frac{Y(0)}{L(0)}\right] \approx\ & gt + \left(1 - e^{-\lambda t}\right)\log(A(0)) \\ & + \left(1 - e^{-\lambda t}\right)\frac{\alpha \log(s_k)}{1-\alpha-\beta} + \left(1 - e^{-\lambda t}\right)\frac{\beta \log(s_h)}{1-\alpha-\beta} \\ & - \left(1 - e^{-\lambda t}\right)\frac{(\alpha+\beta)\log(n+g+\delta)}{1-\alpha-\beta} \\ & - \left(1 - e^{-\lambda t}\right)\log\left[\frac{Y(0)}{L(0)}\right] \end{aligned} \tag{51}$$

Thus, the growth of income per capita is a function of the determinants of the ultimate steady-state equilibrium and the initial level of income.

[26] $e^{-\lambda t} = 1/2$ implies that $t = 34.6$.

To the extent that savings and population growth rates differ across countries and convergence to a steady-state is not complete, the model predicts that there will be differences in per capita income growth rates across countries. MRW test this implication of their model using the same data as before. For all countries, MRW assume that $g = 0.02$ (2% steady-state growth in income per capita) and $\delta = 0.03$ (3% depreciation rate). Rewriting (52) slightly, the equation that MRW estimate is

$$\begin{aligned} \log\left[\frac{Y_i(t)}{L_i(t)}\right] - \log\left[\frac{Y_i(0)}{L_i(0)}\right] &= a + b\log(s_{ki}) + c\log(s_{hi}) \\ &\quad + d\log(n_i + 0.05) + e\log\left[\frac{Y_i(0)}{L_i(0)}\right] + \epsilon_i \end{aligned} \tag{52}$$

where i indexes countries and ϵ_i is the country-specific error term. Results are reported in the second and third columns of Table 9. Standard errors for all coeffecient estimates are in parentheses.

For the broadest sample 98 countries, the results are mixed. On the positive side, all the coefficients have the predicted signs and all except the population growth coefficient are significantly different from zero. Furthermore, the negative coefficient on initial income indicates that there is significant conditional convergence, as is predicted by the theory. On the negative side, the adjusted R^2 of 0.46 is rather low and the implied value of α (0.48) is much higher than it should be ($\alpha \approx 1/3$).

However, for the sample of 22 OECD countries, the results are better. All the coefficients continue the have the predicted signs (although the investment and schooling coefficients are not statistically significant) and the negative coefficient on initial income is even larger, indicating stronger evidence of conditional convergence for OECD countries. Furthermore, the adjusted R^2 of 0.65 is much larger (than in the previous regression for OECD countries) and the implied value of α (0.38) is roughly what it should be ($\alpha \approx 1/3$).

MRW interpret these results as indicating that there are larger departures from steady state for the OECD than for the broader sample of 98 countries. Given that this is the case, it helps explain why the estimated coefficients and R^2's are lower for the OECD in the first specification that does not consider out-of-steady-state dynamics. For the OECD countries, population growth and capital accumulation have not yet had their full impact on standards of living. They attribute the greater departure from steady-state for the OECD

Table 9: Estimation of the Convergence Equation

Dependent variable: log difference GDP per working-age person 1960-85		
Sample:	Non-oil	OECD
Observations:	98	22
Constant	3.04	2.81
	(0.83)	(1.19)
$\log(Y/L(1960))$	-0.289	-0.398
	(0.062)	(0.070)
$\log(I/GDP)$	0.524	0.335
	(0.087)	(0.174)
$\log(n+g+\delta)$	-0.505	-0.844
	(0.288)	(0.334)
log(School)	0.233	0.223
	(0.060)	(0.144)
$\overline{R}^2$	0.46	0.65
Implied α	0.48	0.38
	(0.07)	(0.13)

to World War II, which had a larger negative effect on the OECD than on the rest of the world. Given an estimated value of λ of 0.02, which implies a slow rate of convergence to steady-state, almost half of the departure from steady state in 1945 would have remained by the end of the sample in 1985. Furthermore, the strong evidence of conditional convergence raises questions about the competing Lucas [1988] model where human capital accumulation drives economic growth and conditional convergence does not occur. Mankiw, D. Romer and Weil [1992] conclude that the empirical evidence is generally supportive of their theory. D. Romer [1996] writes,

"Overall, the evidence suggests that a model that maintains the assumption of diminishing returns to capital but adopts a broader definition of capital than traditional physical capital, provides a good first approximation to the cross-country data."

4.1 Criticisms

The MRW paper has attracted a great deal of attention among economists. The thesis that all that needs to be done to "fix" the old growth theory is to make the minor adjustment of adding human capital accumulation certainly is provocative. I will now discuss three important criticisms that have been raised concerning this paper.

First, consider how the regresion results change when R&D investment is included as an explanatory variable. MRW do not allow R&D to play a role in any of their cross-country regressions, but later researchers have. In particular, Lichtenberg [1992] uses the same data as in Mankiw, D. Romer and Weil [1992] but also includes as an explanatory variable, privately funded R&D as a fraction of GDP for each country.[27] Since data on R&D expenditures is less readily available, Lichtenberg is forced to study a smaller sample of 53 countries. Instead of (37), Lichtenberg assumes that the production function for each country i is

$$Y_i(t) = A(t)R_i(t)^{\pi}K_i(t)^{\alpha}H_i(t)^{\beta}L_i(t)^{1-\pi-\alpha-\beta} \tag{53}$$

where $R_i(t)$ is the "research capital" of country i.[28] The intoduction of research capital into the model is the only significant change. Otherwise, the analysis proceeds exactly as in MRW. Supposing that all countries are experiencing steady-state growth, Lichtenberg obtains as a counterpart to (47) and (48)

$$\log\left(\frac{Y_i}{L_i}\right) = c + \frac{\alpha\log(s_{ki}) + \beta\log(s_{hi}) + \pi\log(s_{ri}) - (\alpha+\beta+\pi)\log(n_i+0.05)}{1-\alpha-\beta-\pi} + \epsilon_i \tag{54}$$

---

[27]See also related work by Eaton and Kortum [1993] and Dinopoulos and Thompson [1995].

[28]Lichtenberg's modeling of R&D as generating a rival input in the production process called "research capital" is theoretically unsatisfactory. R&D generates new knowledge, which is an inherently non-rival input unlike any form of capital. As Romer [1994] has emphasized, discoveries differ from other inputs in the sense that many people can use them at the same time. Although the theory is objectionable, Lichtenberg's empirical work is nevertheless interesting because it gives us a rough idea of how important R&D is for explaining cross country differences.

where c is the constant term, s_{ri} is country i's privately funded R&D expenditures as a fraction of GDP and ϵ_i is the country-specific error term. Regression results are reported in Table 10. Lichtenberg uses a different econometric technique for estimating coefficients (nonlinear least-squares estimation) but this is an inconsequential difference. When π is

Table 10: Estimation of the Steady-State Equation

Dependent variable: log GDP per working-age person in 1985		
Observations:	53	53
α	0.282	0.184
	(0.032)	(0.063)
β	0.310	0.321
	(0.021)	(0.043)
π		0.073
		(0.021)

set equal to zero in (54), then the Lichtenberg model becomes the MRW model (with only minor differences). Results are reported for this case in the second column of Table 10. The coefficient estimates are roughly the same as obtained by MRW for their 98 country sample. However, when Lichtenberg allows for non-zero values of π, as is reported in the third column, the coefficient estimates change considerably. There is a big drop in α from 0.282 to 0.184 and π jumps up to 0.073, implying a roughly 7% elasticity of GDP with respect to the privately-funded research capital stock. This later estimate should be viewed as indicating a huge effect of R&D because privately-funded R&D represents only 1.4% of GDP in developed countries and an even smaller percentage in developing countries (the total R&D investment share is only about 0.4% in developing countries). Leaving R&D expenditures out of cross-country regressions appears to be a mistake that seriously biases coefficient estimates.

Lichtenberg also studies convergence issues. As a counterpart to (52) and (53), Lichtenberg obtains

$$\log\left[\frac{Y_i(t)}{L_i(t)}\right] - \log\left[\frac{Y_i(0)}{L_i(0)}\right] \approx \left(1-e^{-\lambda t}\right)\frac{\alpha\log(s_{ki})}{1-\alpha-\beta-\pi} + \left(1-e^{-\lambda t}\right)\frac{\beta\log(s_{hi})}{1-\alpha-\beta-\pi} + \left(1-e^{-\lambda t}\right)\frac{\pi\log(s_{ri})}{1-\alpha-\beta-\pi} - \left(1-e^{-\lambda t}\right)\frac{(\alpha+\beta+\pi)\log(n_i+0.05)}{1-\alpha-\beta-\pi} - \left(1-e^{-\lambda t}\right)\log\left[\frac{Y_i(0)}{L_i(0)}\right] + \epsilon_i \tag{55}$$

Regression results are reported in Table 11. The coefficient estimates in the second column

Table 11: Estimation of the Convergence Equation

Dependent variable: log difference GDP per working-age person 1960-85		
Observations:	53	53
α	0.474	0.354
	(0.047)	(0.086)
β	0.236	0.259
	(0.056)	(0.071)
λ	0.017	0.021
	(0.001)	(0.005)
π		0.066
		(0.026)

are roughly the same as obtained by MRW for their 98 country sample. However, when non-zero values of π are allowed for, as is reported in the third column, coefficient estimates change considerably. There is a big drop in α from 0.474 to a more reasonable 0.354 (remember that α should be roughly one third) and π jumps up to 0.066, implying again a roughly 7% elasticity of GDP with respect to the privately-funded research capital stock. Thus, leaving

R&D expenditures out of the cross-country convergence regressions also appears to be a mistake that seriously biases coefficient estimates.

A second and related criticism of MRW [1992] concerns the assumption that all countries experience the same rate of technological progress, so that the country specific parameter g_i can be replaced by a common parameter g. Grossman and Helpman [1994], for example, argue that this assumption of a common rate of technological progress in all 98 countries over a 25-year period is simply indefensible. The rate at which producers in Japan have acquired new technologies has been markedly different from the rate in Chad, for example. Even restricting attention to OECD countries, Wolff [1992] provides evidence of strikingly different rates of total factor productivity growth over the last 20 years.

This second criticism has econometric implications, as Grossman and Helpman (p.29) point out:

> "If technological progress varies by country and g_i is treated as part of the unobserved error term, then OLS estimates of the (parameter values) will be biased when investment-GDP ratios are correlated with country-specific productivity growth. In particular, if investment rates are high where productivity grows fast, the coefficient on the investment variable will pick up not only the variation in per capita incomes due to differences in countries' tastes for savings, but also part of the variation due to their different experiences with technological progress."

This is a potentially serious problem because there is strong evidence of a positive correlation between total factor productivity growth and the investment ratio.[29] Investment tends to be high when productivity is growing rapidly.

A third criticism that has been raised is that the model fails to explain either rate-of-return to investment differences across countries or international capital flows. This criticism of the Solow [1956] model has been raised by Lucas [1988], and it applies with equal force to the MRW model. The argument goes like this: differentiating (37) with respect to K and then substituting steady state values using (45) and (46), we obtain the steady-state

[29]See Baumol et al. [1989]

marginal product of capital, net of depreciation

$$MPK_i - \delta = \frac{\alpha(n_i + g + \delta)}{s_{ki}} - \delta \tag{56}$$

Thus, the net marginal product of capital in country i depends on that country's population growth rate and savings rate. Because these rates differ substantially across countries, (56) implies that there will be substantial differences in rates of return, with much higher rates of return is poor countries where savings rates are low and population growth rates are high. For example, if one assumes that $\alpha = 1/3$, $\delta = 0.03$, $g = 0.02$, then the data that MRW use implies that the net marginal product of capital is roughly three times as high in the Philippines (a typical poor country) as in Sweden, and twice as high in the Philippines as in the United States.[30] Under the circumstances, it is a bit puzzling why any investment at all takes place in rich countries. In reality, real interest rate differentials are much smaller than this theory predicts and Feldstein and Horioka [1980] have documented that countries with high savings rates have high rates of domestic investment rather than large current account surpluses: capital does not flow from high-saving countries to low-saving countries.

Mankiw, D. Romer and Weil respond that while the model predicts that the marginal product of capital will be high in poor countries, if investors are not optimizing or capital markets are not perfect, then it does not follow that real interest rates will also be high in poor countries. They conjecture that some of the most productive investments in poor countries are in public capital and that the behavior of governments in these countries is not socially optimal. Also the fear of future expropriation may have a big effect on the behavior of potential investors. Williams [1975] found that, from 1956 to 1972, developing country governments nationalized about 19 percent of foreign capital and that compensation averaged about 41 percent of book value. Thus, although there are still puzzles to be resolved, this particular criticism may not be that serious. Furthermore, the opposite side of the coin is that the MRW model does correctly predict the migration incentives of workers, namely that workers want to migrate from poor to rich countries where they earn higher wages.

[30]The net marginal product of capital is 0.043 in Sweden, 0.072 in the United States and 0.148 in the Philippines.

5 Conclusions

My description of the new growth theory has been incomplete; it did not cover all existing approaches nor did it deal with all issues that have been treated in the literature. Admittedly, my choices exhibit are large dose of subjectivity. This type of bias is, however, unavoidable, and I did not intend to survey the literature. I focused my presentation on one important mechanism for sustaining long-run growth, namely, R&D investment by firms and two models (by Grossman and Helpman [1991] and Segerstrom [1996]) that illustrate how growth is sustained. Perhaps the main conclusion that emerges from comparing these models is that slight differences in assumptions about the returns to R&D investment generate significant differences in implications. I have also restricted attention to a limited number of public policy issues to demonstrate the usefulness of viewing economic growth as an R&D-driven phenomenon.

My coverage of human capital accumulation in this paper has been limited to explaining the properties of one model (Mankiw, Romer and Weil [1992]) where the rate of technological change is exogenously determined. One advantage of studying this model and its empirical tests is that one can assess how well models with physical and human capital accumulation alone (no separate knowledge creation activities) can account for cross country differences in income levels and growth rates. Mankiw, Romer and Weil conclude that their model does a good job but as I have discussed in this paper, some potentially serious criticisms have been raised about this model. To date, not much progress has been made on carefully encorporating physical and human capital accumulation into models of R&D-driven growth with many countries, but ultimately, I believe that such models will do a significantly better job of accounting for cross-country differences in income levels and growth rates. I view the "old growth theory" with its emphasis on physical capital accumulation, the "new growth theory" that emphasizes human capital accumulation and the "new growth theory" that emphasizes knowledge creation as being complements rather than substitutes. We will undoubtedly see much work on integrating insights from these branches of the growth literature in the coming years.

References

Aghion, P. and Howitt, P. [1992], "A Model of Growth Through Creative Destruction," *Econometrica,* 60, 323-351.

Baumol, W., Blackman, S., and Wolff, E. [1989], *Productivity and American Leadership: The Long View.* (Cambridge: MIT Press).

Baumol, W. and Blinder, A. [1988], *Economics: Principles and Policy.* (San Diego: Harcourt Brace Jovanovich).

Barro, R. and Sala-I-Martin, X. [1995], *Economic Growth* (New York: McGraw Hill).

Benhabib, J. and Spiegel, M. [1994], "The Role of Human Capital in Economic Development:Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics,* 34, 143-174.

Bils, M. and Klenow, P. [1996], "Does Schooling Cause Growth or the Other Way Around," (working paper, University of Rochester).

Coe, D. and Helpman, E. [1993], "International R&D Spillovers," (Cambridge: National Bureau of Economic Research working paper no. 4444).

Davidson, C. and Segerstrom, P. [1996], "R&D Subsidies and Economic Growth," (working paper, Michigan State Univerity).

Dinopoulos, E. and Segerstrom, P. [1996], "A Schumpeterian Model of Protection and Relative Wages," working paper, University of Florida).

Dinopoulos, E. and Thompson, P. [1995], "Is Endogenous Growth Empirically Relevant?" (working paper, University of Florida).

Eaton, J. and Kortum, S. [1993], "International Technology Diffusion," (working paper, Boston University).

Evenson, R. [1984], "International Invention: Implications for Technology Market Analysis," in *R&D, Patents and Productivity,* ed. Z. Griliches (Chicago: University of Chicago Press).

Feldstein, M. and Horioka, C. [1980], "Domestic Savings and International Capital Flows," *Economic Journal,* 40, 314-29.

Griliches, Z. [1990], "Patent Statistics as Economic Indicators: A Survey," *Journal of Eco-*

nomic Literature, 28, 1661-1707.

Grossman, G. and Helpman, E. [1991], "Quality Ladders in the Theory of Growth," *Review of Economic Studies,* 58, 43-61.

Grossman, G. and Helpman, E. [1994], "Endogenous Innovation in the Theory of Growth," *Journal of Economic Perspectives,* 8, 23-44.

Hoel, P., Port, S. and Stone, C. [1971], *Introduction to Probability Theory* (Boston: Houghton Mifflin).

Jones, C. [1995a], "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics,* 110, 495-525.

Jones, C. [1995b], "R&D-Based Models of Economic Growth," *Journal of Political Economy,* 103, 759-784.

Jones, L. and Manuelli, R. [1990], "A Convex Model of Equilibrium Growth," *Journal of Political Economy,* 98, 1008-1038.

Kortum, S. [1996], "Research and Productivity Growth: Theory and Evidence from Patent Data," National Bureau of Economic Research working paper No. 4646, Cambridge, MA.

Kremer, M. [1993], " Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics,* 108, 681-716.

Lach, S. and Schankerman, M. [1989], "Dynamics of R&D and Investment in the Scientific Sector," *Journal of Political Economy,* 97, 880-904.

Lichtenberg, F. [1992], "R&D Investment and International Productivity Differences," (working paper no. 4161, National Bureau of Economic Research).

Lucas, R. [1988], "On the Mechanics of Economic Development," *Journal of Monetary Economics,* 22, 3-42.

Machlup, F. [1962], *The Production and Distribution of Knowledge in the United States* (Princeton: Princeton University Press).

Mankiw, G., Romer, D. and Weil, D. [1992], "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics,* 107, 407-437.

National Science Board [1993], *Science and Engineering Indicators-1993.* (Washington, D.C.: U.S. Government Printing Office).

Rebelo, S. [1991], "Long Run Policy Analysis and Long Run Growth," *Journal of Political*

Economy, 99, 500-521.

Richardson, D. [1995], "Income Inequality and Trade: How to Think, What to Conclude," *Journal of Economic Perspectives,* 9, 33-55.

Rivera-Batiz, L. and Romer, P. [1991], "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics,* 106, 531-56.

Romer, D. [1996] *Advanced Macroeconomics.* (New York: McGraw Hill).

Romer, P. [1986], "Increasing Returns and Long Run Growth," *Journal of Political Economy,* 94, 1002-1037.

Romer, P. [1990], "Endogenous Technological Change," *Journal of Political Economy,* 98, S71-S102.

Romer, P. [1994], "The Origins of Endogenous Growth," *Journal of Economic Perspectives,* 8, 3-22.

Rosenberg, N. and Nelson, R. [1993], "The Role of University Research in the Advancement of Industrial Technology," (working paper, Stanford University).

Scherer, F. [1980], *Industrial Market Structure and Economic Performance* (Boston: Houghton-Mifflin).

Schumpeter, J. [1942], *Capitalism, Socialism and Democracy* (New York: Harper & Row).

Segerstrom, P. [1991], "Innovation, Imitation and Economic Growth," *Journal of Political Economy,* 99, 807-27.

Segerstrom, P. [1996], "Endogenous Growth without Scale Effects," (working paper, Michigan State University).

Segerstrom, P., Anant, T., and Dinopoulos, E. [1990], "A Schumpeterian Model of the Product Life Cycle," *American Economic Review,* 80, 1077-1092.

Solow, R. [1956], "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics,* 70, 65-94.

Stockman, A. [1996], *Introduction to Economics.* (Fort Worth, TX: Dryden Press).

Summers, R. and Heston, A. [1988], "A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950-85," *Review of Income and Wealth,* 34, 1-26.

Williams, M. [1975], "The Extent and Significance of Nationalization of Foreign-owned Assets in Developing Countries, 1956-1972," *Oxford Economic Papers,* 27, 260-73.

Wolff, E. [1992], "Productivity Growth and Capital Intensity on the Sector and Industry Level: Specialization among OECD Countries, 1970-1988," paper presented at the MERIT Conference on "Convergence and Divergence in Economic Growth and Technical Change," Maastricht.

Young, A. [1995], "Growth Without Scale Effects," National Bureau of Economic Research working paper No. 5211, Cambridge, MA.