6

UNIVERSITÀ DEGLI STUDI DI SIENA
Dipartimento di Economia Politica

COLLANA DEL DIPARTIMENTO
DI ECONOMIA POLITICA

ROBERT M. SOLOW

SIENA LECTURES ON ENDOGENOUS GROWTH THEORY

edited by Serena Sordi

Robert M. Solow

SIENA LECTURES ON ENDOGENOUS GROWTH THEORY

edited by Serena Sordi

Siena, September 1992

Robert M. Solow is Professor of Economics at the Massachusetts Institute of Technology and Salimbeni Professor at the University of Siena for the academic year 1991-92.

PREFACE

When I was invited to give six lectures to Ph.D. students at Siena in the summer of 1992, it seemed to be a good occasion to take stock of the "*New Growth Theory*" of Romer, Lucas and their successors. This book is the result. I am a teacher by profession, and the lectures were intended to teach. Precisely for that reason, they reflect my own peculiar notions about the kind of subject economics is and the way it should be taught.

The reader (or student) will not find here an attempt to build growth theory from first principles or in great generality. I take growth theory to be the way economics tries to think about the mechanisms or interrelationships that have a serious influence on the sort of medium-to-long-run path traced out by the modern industrial capitalist economies we observe, especially after we have averaged out their minor fluctuations. So I focus on a few individual models, each of which aims a spotlight on the particular piece of economic machinery its author believes to be especially important.

I do not for a moment think that this selection of models is in any way complete. For instance, nothing is said about the possible integration of the theory of growth and the theory of short-run macroeconomic equilibrium and/or disequilibrium. I think that is a very important topic. But little has been done, with one or two notable exceptions, and especially little has been done by the *New Growth* theorists whose work has brought such a revival of interest and excitement to the field.

Of course I look at the new models of growth as a veteran of the old models of growth. That is not great strain, because most of what the new models do builds upon the older literature. One of my minor goals is to think through for myself which aspects of the new models represent, theoretically or empirically, substantial improvements on the old ones. Of course I hope also to provide some hints for students who might like to work in this field.

I want to thank the *Dipartimento di Economia Politica* at Siena for the invitation that led to these lectures. They were attended by a small but excellent group of Ph.D. students whose interest and enthusiasm made the lectures more interesting than they might have been. Especially, I am grateful to Dr. Serena Sordi who converted my lecture notes and lectures first into a coherent English text and then into a book in Italian. Her mathematical notes will certainly improve the pedagogical value of the book. But more than that, she has made these lectures better than they really were.

R.M.S.

FIRST LECTURE

I.1. Introduction

I was asked to give a series of lectures on the theory of growth, especially the many recent developments in what is called the *theory of endogenous growth.* My goal is to give you a survey of the *economic* ideas. I will pay minimal attention to questions of technique because the techniques are quite traditional and those of you who have had any kind of introduction to optimization theory or to traditional economic dynamics will find nothing especially new technically.

The version of the theory of growth that I want to cover has already made *two* preliminary decisions which limit its scope very much.

The *first* is that we are going to be talking about completely aggregated models, one-commodity models in effect. All of the "structural" questions are ignored, e.g., questions that have to do with the way in which the proportional importance of different industries is likely to change during a long period, or questions such as the relation between agriculture and industry, or between manufacturing and services.

The *second* decision — and this is true of almost all of growth theory, aggregated or not — is that it ignores what is now called the "*coordination problem*": there is always full employment (or there is always constant unemployment). I will make no attempt — and neither has the profession — to integrate the theory of growth and the theory of economic fluctuations. You discuss the long term and you discuss macroeconomic fluctuations in different series of lectures. Both of those questions deserve attention, but I am not going to discuss either because it would take too much time, as in the case of structural questions, or because not very much is known, as in the case of fluctuations.

I.2. The Standard Neoclassical Model

I am going to begin by discussing the standard neoclassical model. Most of what I want to lecture about are contemporary extensions of that model. But so that we have a platform from which to start, I want to discuss the perfectly standard growth theory of the 1950s and 1960s. I have some things to say which may be unfamiliar to you.

The place to start, remembering that this is a completely aggregated theory, is to express output (Y) as a function of three things: (i) the stock of capital (K), which consists of an accumulated stock of the single output, (ii) the current volume of employment (N), and (iii) time itself, so as to suggest that the relationship between output, capital input and labor input may change through time:

$$Y = F(K,N,t).$$

In addition, this output is divided into two components, one of which is consumed and the other added to the stock of capital:

$$Y = Nc + \dot{K},$$

where $\dot{K}$ is the time derivative of the stock of capital. It is convenient to use the symbol c to stand for consumption *per capita* so that aggregate consumption is Nc.

What should be said about the determination of the total amount of consumption?

There are two directions that the literature has taken in the 1950s-1960s and up until the present time.

One of these I will call "*behavioristic*". The idea here is to assume any kind of empirically plausible consumption function. There is some sort of consumption function, or saving function, or perhaps some different formulation, which is justified in the way we always justify consumption functions: it seems to make sense, it fits the data, etc. So let us start with a consumption function which must be a function of the variables which appear in the model:

$$c(K,N,t).$$

One simply assumes this. Then, the analysis of this kind of model reduces to study the following differential equation:

$$\dot{K} = F(K,N,t) - Nc(K,N,t).$$

The usual assumption is that the level of employment, which is defined essentially to be the same as or proportional to the level of population, is growing exponentially:

$$N = N_0 e^{nt}.$$

The "coordination problem" disappears here because it is assumed that the level of employment is always equal to the size of the labor force. Much of macroeconomics is about — if we replace the argument here by employment (L) — why and when L might differ from N. But the tradition in macroeconomics is to separate the study of this from the study of growth. I do not think that it is a good tradition, but it is the tradition and very little has been done outside that tradition.

The "*behavioristic*" version of aggregate growth theory just boils down to thinking about $F(K,N,t)$, thinking about $c(K,N,t)$, and studying the solution to the above differential equation which will tell us how the economy evolves from any initial condition.

The *second* branch of growth theory we can think of as "*optimizing*" theory and the assumption then is that the economy we are studying behaves *as if* it were inhabited by a single, immortal household and that household is concerned with behaving optimally. The problem which is traditionally posed to the household is that it should choose the path of *per capita* consumption so as to maximize:

$$\int_0^{\infty} e^{-\rho t} u(c(t)) N(t) dt,$$

i.e., to maximize the integral of instantaneous utility, where ρ is the discount rate. The utility function is conventionally defined as a *per capita* function and usually one multiplies it by the size of the population (N). The household would like to maximize the discounted and weighted sum of utilities, where the weights are the size of the population. The natural interpretation of this is that the agent is a peasant household, an isolated peasant family. The individual may even die but the family goes on for ever, with consistent preferences.

This maximization is subject to a constraint which essentially comes from the technology, namely the constraint that:

$$N(t)c(t) + \dot{K} = F(K,N,t).$$

The "*optimizing*" version of one-sector growth theory says that the economy behaves *as if* it was solving this problem. You see where the "coordination problem" has vanished here: the production side of this economy just performs what is best for the household. Under favorable assumptions, the unique solution to this problem is also the unique competitive equilibrium for this economy.

I say competitive equilibrium in a very special meaning because everything happens over time: either it is the competitive equilibrium for this economy with what I will call Arrow-Debreu markets, meaning that markets for every instant of time are available at the beginning of the problem, or else it would do just as well if everyone in this economy had infinite perfect foresight. There is a sense in which the optimizing version of this model could be decentralized into a perfectly competitive economy, but it has to be a perfectly competitive economy under the most favorable assumptions, the sort of assumptions that rule out coordination failures, and in particular the Arrow-Debreu assumption that all the markets are there at the very beginning of time.

That in a way is the story: in one version one starts with the technology given and studies differential equations, in the other version one studies the solution to the above optimization problem. You will see that it hardly matters for the steady states which one of these two branches one chooses. The theory looks very much the same.

Before I go on, I want to discuss two assumptions about the production function that are usually made in the literature and that are usually misunderstood.

The first assumption that I want to study is that the production function $F(K,N,t)$ can be written as:

$$F(K,A(t)N),$$

with the usual further assumption that $A(t)$ is exponential:

$$A(t) = e^{at}.$$

If you think of the dependence of F on t as representing technological progress — the fact that techniques of production improve all the time — then this formulation is called *labor-augmenting* technological progress. Essentially all of one-sector growth theory — both the old growth theory of 1950s and 1960s and the new growth theory of the second half of the 1980s — is carried on under an assumption like this. It appears to be a very arbitrary assumption. I want to convince you that it is in a sense arbitrary, but that the arbitrariness does not lie in the choice of this functional form; the arbitrariness lies elsewhere and has nothing to do with it.

In order to show you what I mean I will simply work out a kind of example.

Suppose we imagine that F has constant returns to scale, it is homogenous of degree one in the first argument and in the second argument, and suppose, instead of the labor-augmenting case, a more general form:

$$F(e^{bt}K, e^{at}N),$$

namely, suppose we try labor- and capital-augmenting technological progress.

I want to show you under what circumstances one cannot deal with this case, i.e., under what circumstances b must be zero. If b is zero, then we are back to the labor-augmenting case. Then you will see what explains the preference for insisting on this rather special form of technological progress.

It will be adequate if I consider for this case what happens if investment is simply proportional to output. Thus:

$$\dot{K} = sF(e^{bt}K, e^{at}N).$$

Suppose we want to look for exponential steady states, i.e., for a situation in which:

$$K = K_0 e^{gt}.$$

Then of course:

$$\dot{K} = gK_0 e^{gt}.$$

So, what we must have in the steady state is:

$$\begin{aligned} gK_0e^{gt} &= sF(e^{bt}K, e^{at}N) = \\ &= sF(e^{(b+g)t}K_0, e^{(a+n)t}N_0) = \\ &= se^{(a+n)t}N_0F\left(\frac{K_0}{N_0}e^{(b+g-a-n)t}, 1\right). \end{aligned}$$

If the differential equation is to have a solution that traces out an exponential steady state, then this condition must hold.

The left-hand side of this condition is an exponential growing at the rate g and the right-hand side is the product of an exponential growing at the rate $(a + n)$ and of the function F evaluated at $(K,N) = ((K_0/N_0)e^{(b+g-a-n)t}, 1)$. This last expression then has to be an exponential because any other function of time multiplied by an exponential would not get you an exponential. There are only two ways in which this condition can hold.

First of all, it could be that b is zero, and g equals $(a + n)$:

$$gK_0e^{gt} = sN_0e^{gt}F\left(\frac{K_0}{N_0},1\right) = sF(K_0e^{gt},N_0e^{gt}),$$

and this is exactly the labor-augmenting case.

Secondly, it could be that the function F is actually exponential as well:

$$F(x,1) = x^c.$$

If F is an exponential, the only way that this function of an exponential can be an exponential itself is Cobb-Douglas with:

$$c = \frac{g-a-n}{b+g-a-n},$$

so:

$$c(b + g - a - n) = g - a - n,$$

or

$$g = a + n + \frac{bc}{1-c}.$$

In this case:

$$(e^{bt}K)^c(e^{at}N)^{1-c}$$

can be written as

$$K^c(e^{[a+bc/(1-c)]t}N)^{1-c}$$

which is the labor-augmenting case again.

Thus, the devotion of this kind of theory to labor-augmenting technological progress exactly corresponds to our interest in exponential steady states. If we were to lose interest in exponential steady states, then there would be no need to have this assumption about the nature of technological progress. We could simply choose anything that you like as consumption function and solve the differential equation. Unless F were this particular form there would never be an exponential steady state. The moral of this part of the story is that labor-augmenting technological progress is not a special assumption which is needed for this kind of theory to work out; it is a special assumption which is needed so that we poor people can talk about exponential steady states.

The second assumption that I want to analyze it is important because again is much misunderstood.

It is often thought that the decisive innovation of the new growth theory with endogenous rates of growth is that it assumes increasing returns to scale. I want to show you that that is not so, i.e., that increasing returns to scale does not help in getting endogenous rates of growth.

Let us try to give an example of that.

In doing the previous argument, I assumed that F, as a function of two variables, had constant returns to scale. Now what I want to do is to assume that it has increasing returns to scale. I want to show you just how much change that makes. It turns out to be very little, and it does not open the way to endogenous determination of the rate of growth. That is a misconception.

Let $F(x,y)$ be homogeneous of degree one, and let the production function be $F(K,AN^h)$, where $h > 1$. Notice that this *does* have increasing returns to scale in K and N. To see that, multiply K and N by a number $\lambda > 1$. We then have:

$$F(\lambda K,(A\lambda N)^h) = \lambda F(K,\lambda^{h-1}(AN)^h) > \lambda F(K,(AN)^h),$$

because $\lambda^{h-1} > 1$.

We see that this is a formulation which gives us increasing returns to scale. Now I want to do exactly the same exercise that I did over before. Suppose that we look at an exponential steady state for this model:

$$K = e^{gt}.$$

So we would have:

$$ge^{gt} = sF(e^{gt}, e^{(a+n)ht}).$$

This can be a steady state in one and only one circumstance, that is if:

$$g = (a + n)h.$$

So if there are increasing returns to scale, if $h > 1$, then the only conceivable steady state growth rate is $h(a + n)$. When $h = 1$, we are back to constant returns to scale. But if there are increasing returns to scale in the only form that allows for exponential steady state, then you get a rate of growth which is exogenous.

You get something very paradoxical out of these two equations that I want to point out. $(g - n)$, which is the rate of growth of output *per capita*, would be equal to:

$$g - n = (a + n)h - n = \\ = ha + (h - 1)n,$$

and that would mean that the rate of growth of productivity is faster in an economy which has a faster rate of population growth because $(h - 1)$ is certainly positive. This is not a very promising assumption. There is no evidence at all, and hardly anybody in the world believes, that faster population growth implies faster productivity growth. In any case I am not much interested in this point, and I am not about to develop it further.

What I *did* want to emphasize is this: when you allow increasing returns to scale in a model of this kind in such a way as to allow an exponential steady state to occur, that does not achieve for you in any way an endogenizing of the growth rate. The conclusion is that increasing returns to scale is *not* the key to *endogenous growth*. One of the economic ideas that I want you to get in the course of these lectures is that the extra thing that you need to endogenize the growth rates is usually very strong, much stronger and much more critical an assumption than increasing returns to scale.

I want now to go on and remind you of how the "*optimizing*" version of the neoclassical growth model works. Then, we will be able to use that as a basis for moving on to the new growth theory.

I am going to make all of the usual simplifying assumptions such as, for instance, that the utility function defined on *per capita* consumption is:

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma},$$

where $\sigma > 0$, and where $(1/\sigma)$ is the intertemporal elasticity of substitution which measures how easily the consumer or the household can substitute consumption at different points of time. Thus, $\sigma = 1$ — in which case the function reduces to the logarithmic utility function — is a sort of central case. When $\sigma > 1$, then consumptions at different time are poor substitutes for one another; when $0 < \sigma < 1$ the elasticity of substitution is bigger than one and the consumer finds it much easier to trade consumption now for consumption later. Obviously, when $\sigma = 0$ we have the linear case.

The job of the consumer is to maximize by choice of the consumption path the following integral:

$$\int_0^{\infty} e^{-\rho t}\left(\frac{c(t)^{1-\sigma} - 1}{1-\sigma}\right)N(t)dt,$$

where ρ, which is positive, measures the utility discount rate.

This maximization has to occur subject to the following constraint:

$$N(t)c(t) + \dot{K} = F(K, A(t)N). \tag{1}$$

Always for this discussion I will choose:

$$N(t) = e^{\lambda t},$$

and

$$A(t) = e^{\mu t},$$

i.e., the number of people in this peasant-family is growing exponentially at the rate λ and the rate of labor-augmenting technological progress is μ, where λ and μ are both positive.

The "*optimizing*" version of the theory says that when you look out at a growing economy what it is doing is tracing out a path that solves this optimization problem.

The standard technique for solving problems like this is to form what is called the current-value Hamiltonian, which in this case consists of:

$$H = \left(\frac{c^{1-\sigma} - 1}{1-\sigma}\right)N + p[F(K, AN) - Nc].$$

The first term on the right-hand side is the current flow of utility, and the second term is net investment multiplied by the shadow price of net investment. So there is a sense in which the current-value Hamiltonian is the net national product at each instant of time. It is the consumption measured in utility units plus investment measured in shadow price and you will see that this shadow price really does convert investment into utility.

To solve the problem, there are really only three things that we have to remember. The first is that the Hamiltonian must be maximized with respect to c at each instant of time, and that implies the following first order condition:

(2) $$p(t) = c(t)^{-\sigma}.$$

The term on the right-hand side is the marginal utility of consumption, so that equation (2) implies that the shadow price of investment at each instant of time must be equal to the marginal utility of consumption at that instant of time. The peasant household at each instant has an output and can allocate it between consumption and investment. If it is going to do the best it can over time, it will allocate the output so that the marginal gain from putting a little more into consumption is just equal to the marginal loss from taking that little bit away from investment. The gain from a little more of consumption is the marginal utility of consumption. The loss from reducing investment by a little bit is the shadow price. Thus, equation (2) must hold at every instant of time.

The next equation is an interesting mathematical result which is the so-called *co-state equation*. This equation tells us something more about the shadow price p. It says that:

(3) $$\dot{p} = \rho p - \frac{\partial H}{\partial K} = p\left(\rho - \frac{\partial F}{\partial K}\right) = p(\rho - F_k),$$

so that:

$$\rho = \frac{\dot{p}}{p} + F_k,$$

and this is the *Fisher equation*. It says that the sum of the marginal product of capital plus the capital gain per unit of capital must equal the pure rate of time preference. This is another, explicitly intertemporal, necessary condition for solving the problem. If it does not hold at any instant, some intertemporal substitution could improve welfare.

(3) is a differential equation in p and (1) is a differential equation in K, but c appears in (1). We could eliminate c using (2) and then, replacing N and A by their known forms, (1) and (3) are two ordinary differential equations in p and K.

There is only one initial condition:

$$K(0) = K_0,$$

so that, as a consequence, there is a one-parameter family of solutions.

To know which of those solutions is the right solution we need one more condition that must be satisfied and that is the so-called *transversality condition*:

$$\lim_{t\to\infty} e^{-\rho t} p(t)K(t) = 0.$$

In a well-behaved problem like this, there is only one of this one-parameter family of solutions that satisfies this condition. In this way we find a solution to the problem.

Now, I want to go over to the Cobb-Douglas assumption, and I want to think about the steady states:

$$F(K,AN) = K^{\beta}(AN)^{1-\beta}.$$

In this special case equation (3) becomes:

$$\frac{\dot{p}}{p} = [\rho - \beta K^{\beta-1}(AN)^{1-\beta}].$$

I am now going to look only at exponential steady states, when p, K, and c are exponential at constant rates of growth.

Let us start by calling the rate of growth of consumption per head γ:

$$\hat{c} = \frac{1}{c}\frac{dc}{dt} = \gamma.$$

I will eventually find that γ can be evaluated in a very simple way in terms of the parameters of the model.

Equation (2) says that:

$$\dot{p} = -\sigma c^{-\sigma-1}\dot{c},$$

i.e.:

$$\frac{\dot{p}}{c^{-\sigma}} = \frac{\dot{p}}{p} = \hat{p} = -\sigma\left(\frac{\dot{c}}{c}\right) = -\sigma\hat{c} = -\sigma\gamma,$$

so that, from equation (3):

$$F_k = \beta K^{\beta-1}(AN)^{1-\beta} = \rho + \sigma\gamma,$$

ρ, σ and γ are constants and so what this tells us is that in any steady state the marginal product of capital will be a constant equal to the time discount rate plus σ times the rate of growth of consumption per head.

With a Cobb-Douglas production function, we have:

$$\frac{F}{K} = K^{\beta-1}(AN)^{1-\beta} = \left(\frac{1}{\beta}\right)F_k,$$

so that, if in the steady state the marginal product of capital is a constant, the average product of capital is that same constant divided by β.

Now, let us have a look at equation (1) and divide both sides by K. We obtain:

$$\frac{Nc}{K} + \frac{\dot{K}}{K} = \frac{F}{K},$$

so (Nc/K) is constant in steady state. Since (Nc/K) is a constant, its time derivative must be zero and that tells us that:

$$\hat{N} + \hat{c} - \hat{K} = \lambda + \gamma - \hat{K} = 0;$$

thus:

$$\hat{K} = \lambda + \gamma,$$

is the steady state rate of growth of K.

We have seen that F_k is constant in steady state:

$$F_k = \beta K^{\beta-1}(AN)^{1-\beta}.$$

Thus:

$$\begin{aligned}\dot{F}_k &= \beta(\beta-1)K^{\beta-2}\dot{K}(AN)^{1-\beta} + \beta K^{\beta-1}(1-\beta)(AN)^{-\beta}(\dot{A}N + A\dot{N}) = \\ &= \beta K^{\beta-1}(AN)^{1-\beta}\left\{(\beta-1)\frac{\dot{K}}{K} + (1-\beta)\left(\frac{\dot{A}N + A\dot{N}}{AN}\right)\right\} = \\ &= 0,\end{aligned}$$

i.e.:

$$\begin{aligned}\frac{\dot{F}_k}{F_k} &= (\beta-1)\frac{\dot{K}}{K} + (1-\beta)\left(\frac{\dot{A}}{A} + \frac{\dot{N}}{N}\right) = \\ &= 0,\end{aligned}$$

i.e.:

$$(\beta - 1)(\lambda + \gamma) + (1 - \beta)(\mu + \lambda) = 0,$$

so finally:

$$\gamma = \mu.$$

One other characteristic of the steady state is very important. We can ask what is the investment quota in the optimal steady state.

We have:

$$\begin{aligned}\frac{\dot{K}}{\dot{K} + Nc} &= \frac{\dot{K}/K}{(\dot{K}/K) + (Nc/K)} = \\ &= \frac{\lambda + \gamma}{(\dot{K}/K) + (F/K) - (\dot{K}/K)} = \\ &= \frac{\lambda + \gamma}{(1/\beta)F_k} = \\ &= \frac{\beta(\lambda + \gamma)}{\rho + \sigma\gamma} = \frac{\beta(\lambda + \mu)}{\rho + \sigma\mu}.\end{aligned}$$

Now we have a complete description of all that is interesting about the steady state. Consumption per head is growing at the rate γ, the rate of labor-augmenting technological progress. The stock of capital is growing at a rate equal to the sum of the rate at which the population is growing and the rate of capital-augmenting technological progress. Output is growing at the same rate.

I can divide the parameters of this model into three classes:

(1) technological parameters: β, μ;
(2) demographic parameters: λ;
(3) preference parameters: ρ, σ.

It is an interesting observation that the real growth rates — the growth rates of c, Y and K — depend only on the technological and demographic parameters, but not on the preference parameters. *That is what is meant by saying that this is a model of exogenous growth.* On the other hand, it is interesting that the investment-output ratio, the asymptotic steady state investment-output ratio, depends on all the parameters, and it depends on all the parameters in a reasonable way. The faster the population is growing, the more investment there will be. The bigger the rate of time discount, the less investment there will be. If the population preferences favor current consumption rather than future consumption, then naturally there will be less investment. If the elasticity of substitution between present and future consumption ($1/\sigma$) is small, then there will be relatively little investment because if future consumption is a very poor substitute for present consumption, the optimizing population will not be inclined to save and invest very much because what it gets from that is future consumption.

The one remaining thing I have to pick up is the tranversality condition because that tells something interesting. The tranversality condition was:

$$\lim_{t\to+\infty} e^{-\rho t} p(t)K(t) = 0.$$

The term $p(t)K(t)$ is growing at the rate $(\lambda + \mu - \sigma\mu)$. Thus, in order for the tranversality condition to be true, ρ must be bigger than $(\lambda + \mu - \sigma\mu)$, i.e.:

$$\rho > \lambda + \mu - \sigma\mu,$$

i.e.:

$$\rho + \sigma\mu > \lambda + \mu.$$

This is the so-called *modified golden rule*. Notice that it implies that the investment quota is less than β (the *unmodified golden rule*).

I.3. Conclusions

I have six conclusions.

The first conclusion is that the rates of growth of consumption per head, of output per head and of capital per head are all equal to the exogenously given rate of labor-augmenting technological progress μ. So the steady state rates of growth are exogenous. They are not determined within the model.

Conclusion number two, which I have not discussed at all but which I hope you will remember from the literature, is that the steady state is approached asymptotically by any optimal path from any initial conditions: starting from arbitrary initial conditions the solution to the optimization problem converges to the saddle point.

The third conclusion is that in the steady state the investment-output ratio is a constant depending in a normal way on technology (β,μ), on demography (λ), and on tastes (ρ, σ).

The fourth conclusion is that we lose very little, from the steady state point of view, by adopting the "behavioristic" assumption of a constant investment-output ratio. For any reasonable s, there will be values of ρ and σ that make it "optimal". There is an important question for economists here. Should we regard ρ and σ as "deep" parameters and s as a "superficial" parameter? The answer to that question is 'yes' if you think the real economy is really tracing out the optimizing path of the single and immortal peasant household. If you do not think that, then s is just about as deep a parameter as ρ and σ.

The fifth conclusion I want you to remember is that the transversality condition tells us something about the "Modified Golden Rule", that the optimal s is less than the elasticity of output with respect to capital.

The last conclusion is that remembering that the growth rates do not depend on the taste parameters at all and therefore do not depend on s, we know that as long as s is less than this quantity, a one time parametric increase in s does not change the growth rate but it moves the economy to a higher utility path. A higher value of s or a lower value of ρ or a lower value of σ — it depends on how you want to look at it — will change the steady state path to a higher level as long as we do not exceed the "Modified Golden Rule", but the rate of growth will be the same.

Those are the conclusions that I want to come to. What I am going to do in the next lecture is to begin by discussing in general terms one of the reasons for being unhappy with this model. I want to describe the reasons why within the last five or six years the economic profession has felt that it is important to extend this model in various important ways. Then I want to give you a first example of one of the ways of extending the model that has appeared in the literature. I will choose first Lucas's 1988 paper. What I want to do is to discuss the kind of changes that are made in this model. I want to point out to you that there is a very strong assumption which differentiates Lucas's model from the model we have analyzed today.

SECOND LECTURE

II.1 Introduction

There are several reasons for wanting to extend the standard neoclassical growth model. One set of reasons has to do with saying something about structural problems, i.e., about different sectors. Another reason is to say something about "coordination problems" to allow for what we normally think of as the Keynesian side.

I am not going to discuss that at all, but there two other classes of reasons why one might want to extend the theory.

The first is that it is intellectually unsatisfactory to have the growth rate exogenous. The actual long-run growth rate of an economy is a very important characteristic and to say that is exogenous is not satisfactory. One has to keep in mind that some things *are* exogenous after all. You should also keep in mind that everyone, so to speak, has always known that there is an endogenous side of technological progress. Part of the growth of technology is endogenous. But unless you have a good theory, a reasonable and productive theory of endogenous technological progress, a theory of innovations in other words, it is not worthwhile spending any time on it. We also take λ, the rate of population growth, as exogenous. We all know that population growth is partially endogenous, that has been known since Malthus and no doubt before that. But it would be pointless for me or for Lucas or Domar or anyone to say that the rate of population growth is endogenous unless I have something to say about it. Not having something to say about it, or not having something very interesting or new to say about it, one can just take it as given. In fact in the 1950s I and others did have a little bit to say about the rate of population growth. We had a certain picture of the way the rate of population growth depends on the level of consumption per head. And we also know that paying attention to that possibility can lead to simple models of "poverty trap" and things of that sort. Remember also that there is no problem, in principle, to make a model of endogenous technological progress. The difficulty is to make a good and interesting model.

There are roughly three ways in which growth theory has tried in the last seven or eight years to get beyond a theory in which the growth rate of consumption, and of all the *per capita* quantities, is exogenous. I am going to try to give you examples of some of them.

One device is to study the endogenous accumulation of human capital and I am going to discuss with you in detail today an example of that. You should keep in mind that to add to a growth model an endogenous model of the accumulation of human capital does not guarantee that the growth rate becomes endogenous. I will give you examples of both kinds of neoclassical sort of growth model extended by a model of the accumulation of human capital in which that addition *does* make the rate of growth endogenous and in which it *does not* make the growth rate endogenous. All it depends on how you do it, and one can be skeptical of certain ways of doing it.

The second way of proceeding is actually to have a theory of innovations. In other words, to make μ endogenous by a theory of research, of development, or something of that sort.

There is a third device that I want to mention as well and give you an example, and that is to drop one or more of the standard assumptions of the neoclassical growth model. The one that is usually dropped is diminishing returns to capital. As you will see, without diminishing returns to capital one is back to Domar (1946, 1957), actually. It is rather amazing. I will try to give you an example of this, namely, of the fact that the modern literature is in part just a very complicated way of disguising the fact that it is going back to Domar, and, as with Domar, the rate of growth becomes endogenous.

There is another reason why one might want to change the standard neoclassical model. If that model has some implications that are clearly empirically false, then one would say that we better change this theory. There has been an explosion of work on making international comparisons of growth rates in order to compare the implications of the standard growth model with data.

The reason for this explosion of work is primarily the availability of valid data, the Summers-Heston data. Summers and Heston (1984, 1988, 1991) have produced as best they can comparable National Accounts for something like one hundred and twenty different countries. There are a lot of cross-section statistical studies of determinants of growth rates across countries. Most of that discussion is about convergence or divergence: is it true that all of the nations of the world appear to be convergent to a common rate of growth? Also a lot of work is done about levels: are the levels of income per person in the different countries of the world converging to a common level?

The conclusion from all of these cross-section studies is that the results are not robust. The results can be changed by minor changes in assumptions, by minor changes in the use of data, by minor changes in the time-period for which one operates. Later on we will

go over one or two of those statistical studies just to see how the same data can give apparently different conclusions.

What I want to do now is to describe to you and discuss a slightly generalized version of Lucas's original attempt to get beyond the standard neoclassical model. This has certainly been one of the important pieces of the literature and it is very well done. One of the things I want to show you is a rather surprising result about it, that I heard from two young Italian economists (Paolo De Santis and Giuseppe Moscarini) in Rome not very long ago. For that reason, I am going to analyze a slightly generalized version of the Lucas model. I would rather have done Lucas's own model first and then the generalization, but I think I may not have time to do them both and so it is more important for me to discuss the generalized version with you.

II.2. A Slightly Generalized Lucas Model

The general structure of Lucas's 1988 model is like that of the standard neo-classical model, in the "optimizing" version, i.e., the path of the economy is obtained by maximizing a utility integral, exactly as in the standard neoclassical model.

The economy does what is necessary to maximize:

$$\int_0^{\infty} e^{-\rho t} N(t)\left[\frac{c(t)^{1-\sigma}+al(t)^{1-\sigma}}{1-\sigma}\right]dt,$$

by choice of c, l, and u.

The difference occurs in the utility function. As with the neoclassical model I use a constant elasticity of substitution form, but instead of just this I add a little constant a times another variable called l, which stands for leisure. The constant a is just a weight. *Lucas sets* $a = 0$ so that the second term does not appear. There is no leisure in the Lucas model where — and I will come to this in more detail later on — each member of the population or of the labor force is endowed with one unit of time per unit of time and uses all of that in one of two activities: either working, which is called u in Lucas's paper and I will stick with that notation, or studying. There is no other use of time. Any time that is not spent working is spent accumulating human capital. Or one can be a little more general than that, and say that Lucas assumes that the amount of leisure is fixed exogenously and that there is no choice about it.

What I want to do is to alter that model so that the one unit of time that each member of the population has can be used either for working, or for leisure or for studying. In other words, the individual has the choice to allocate time also to leisure. It turns out that that makes a significant difference in the Lucas model. It is really quite amazing and I was very surprised.

The above integral is what the economy is maximizing and now I want to see what are the constraints.

The first is the standard production function according to which aggregate consumption plus net investment is equal to the quantity produced using a Cobb-Douglas technology:

$$N(t)c(t) + \dot{K}(t) = K(t)^{\beta}[u(t)H(t)]^{1-\beta}\overline{H}(t)^{\gamma}. \tag{1}$$

The labor input consists of u, the fraction of time spent working, times H, where H is the input of labor measured in efficiency units. In this way account is taken of the accumulation of human capital. That is the basis of the model.

Lucas also says that the accumulation of human capital has an external effect as well. If other people have accumulated human capital I will be more productive for any given amount of human capital that I have accumulated. For this reason, he adds an external effect ($\overline{H}$). I will put a bar over this to indicate that for an individual maximizer this is to be regarded as given. For a social planner H would have the exponent $(1 - \beta + \gamma)$ because a social planner would take into account that the accumulation of human capital increases output not only directly, but also through the externality. I am going to produce not the planner solution but the competitive equilibrium solution, and for the competitive equilibrium solution $\overline{H}$ will be regarded as a parameter at each time, as independent of each individual decision.

Notice that the individual agent in the economy is looking at constant returns to scale, at $(\beta + 1 - \beta)$, but the social planner would be looking at increasing returns to scale, at $(\beta + 1 - \beta + \gamma)$. One of the lessons that you must learn, and Lucas says so in the paper, is that γ is not important for his results. As I pointed out at the very beginning of the last lecture, increasing returns to scale is not the secret of anything in a growth model. If γ were zero, so that the external effect did not exist, one would still get the feeling, i.e., the atmosphere of Lucas's model would still be the same.

There is another constraint because, since H, the human capital, appears in equation (1), there must be a model of the accumulation of human capital. The accumulation of human capital goes according to the following rule:

$$\dot{H} = \delta H[1 - l(t) - u(t)]. \tag{2}$$

This is the differential equation that governs the accumulation of human capital and is an extraordinary relationship.

When you think about it, you can see that almost here, in this one line, Lucas is assuming the endogenousness of growth. Also, I think this is very far from a plausible relationship. Let me point out you first that if you think that (2) is a production function for new human capital, and if you think of the inputs as being already accumulated human capital and studying time, then this production function is homogeneous of degree two. It has very strong increasing returns to scale, and constant returns to H itself. If that were not so, if this H were raised to a power less than one, then the Lucas model would not generate *endogenous growth*. The role of studying time is less important. He knows that and in effect he says so, but people seems to have forgotten it.

To see that practically *endogenous growth* has been assumed, just recognize that if u and l are any constant values adding up to less than one, then the growth rate of human capital is already δ times that constant. So simply by changing the constant value of u and l you change the growth rate of human capital. And the growth rate of output is roughly $(1 - \beta)$ times the rate of growth of human capital, so any endogenous decision to change u and l, for example to reduce them a little bit, will increase the growth rate of H and therefore will increase the growth rate of output. The endogenousness comes from saying that everyone will agree that the allocation of time is endogenous, and if the allocation of time is enough to change the growth rates of the factors of production, then of course it will change the growth rate of output. There is nothing complicated or deep about this, it is just as simple as that. Nevertheless you will see something rather strange happening here.

Now, having written down the problem, I want to proceed just as we did in the standard neoclassical model. I start by writing down the current-value Hamiltonian; then we will look at the first order conditions, and finally we will go on and analyze what the rates of growths are. It is straightforward as that.

The current-value Hamiltonian (V) is:

$$V = N(t)\left[\frac{c(t)^{1-\sigma} + al(t)^{1-\sigma}}{1-\sigma}\right] + \\ + p(t)\left\{K(t)^{\beta}[u(t)H(t)]^{1-\beta}\overline{H}(t)^{\gamma} - N(t)c(t)\right\} + q(t)[1 - l(t) - u(t)]H(t),$$

where we had to introduce a second shadow price, $q(t)$, i.e., the shadow price or co-state variable for human capital. As always, the current-value Hamiltonian is a kind of net national product in utility terms.

Now we can do the optimization. First of all, the Hamiltonian has to be maximized instantaneously with respect to $c(t)$, $l(t)$, and $u(t)$. Output has to be allocated between consumption and investment and time has to be allocated between employment, leisure and accumulation of human capital. I get three immediate first order conditions.

The first one, (3a), is exactly what it was before in the standard neoclassical growth model:

$$c^{-\sigma} = p. \tag{3a}$$

Now I differentiate V with respect to leisure and what I get is:

$$Nal^{-\sigma} = q\delta H. \tag{3b}$$

Finally, I have to maximize the current-value Hamiltonian with respect to u, the working time. I get:

$$p(1-\beta)K^{\beta}H^{1-\beta}\overline{H}^{\gamma}u^{-\beta} = q\delta H. \tag{3c}$$

There is an economic meaning to each of these conditions. Since output can be allocated either to consumption or to investment, (3a) says that the marginal utility of consumption must be equal to the value of the marginal utility of net investment which is the shadow price. Since time can be allocated between leisure and work or leisure and studying, or work and studying there are two margins which have to be equated. The marginal value of time consumed as leisure, which is the marginal utility of leisure, must be equal to the marginal value of time spent studying. So (3b) takes care of the leisure-studying margin. The value of the marginal unit of time devoted to study must just be equal to the value of the marginal unit of time devoted to production, and that is condition (3c).

Then we have two more equations which are the co-state equations:

$$\dot{p} = \rho p - \frac{\partial \text{Hamiltonian}}{\partial K} = \rho p - p\beta K^{\beta-1}(uH)^{1-\beta}H^{\gamma}, \tag{4a}$$

$$\dot{q} = \rho q - \frac{\partial \text{Hamiltonian}}{\partial H} = \rho q - p(1-\beta)K^{\beta}u^{1-\beta}H^{\gamma-\beta} - q\delta(1-l-u). \tag{4b}$$

At this stage, I can take the bar off of the Hamiltonian. It is in doing the maximization that one must take account of the fact that $\overline{H}$ is treated as exogenous.

Finally, we have the two transversality conditions:

$$\lim_{t\to\infty} e^{-\rho t}p(t)K(t) = \lim_{t\to\infty} e^{-\rho t}q(t)H(t) = 0.$$

If that were not so, there would be a tendency to postpone consumption for ever.

The next thing I want to do is to proceed as we did with the neoclassical model and to say that these equations characterize the path of the economy. If I believe that we are looking at an economy which behaves *as if* it were maximizing the current-value Hamiltonian subject to those constraints, then the time path followed by the economy would be determined by these equations. There is a unique solution and in a way that is the end of the problem. But I like to know more about the paths and in particular I am going to do just what we did in the standard neoclassical model and look for the steady states.

The steady states are situations in which things like K, c and H have constant rates of growth. Also u and l have constant rates of growth, but those rates of growth will happen to be zero because if l and u grow exponentially obviously very soon they will exceed one. The only possible steady state growth rates for l and u are zero. They could be negative, i.e., they could be decaying toward zero, but that would be very strange: work and leisure go to zero and all time is spent studying! I am going to take the case where u and l in the steady state are constant and that is Lucas's solution too. He does not have an l but he makes u a constant in the steady state.

Now, for some temporary notation, let me call the rate of growth of N λ:

$$\hat{N} = \lambda.$$

This is a parameter. We are supposed to know what the rate of growth of the population is. As I mentioned earlier, if you are interested in applying a model like this to India in the early part of the Twentieth century or perhaps even now you might not want to accept the idea that the population growth rate is constant. You might wish to have an endogenous determination of it. But for Italy or France or US, or possibly even India today, you may be safe to treat population growth as a parameter, which may change from time to time.

Just for notation, I am going to call the rate of growth of consumption per head by the Greek letter χ, the rate of growth of the stock of capital by ξ and the rate of growth of the stock of human capital by υ:

$$\hat{c} = \chi,$$
$$\hat{K} = \xi,$$
$$\hat{H} = \upsilon.$$

Remember that these are all unknowns. We intend to find out what these numbers are in terms of known parameters and functions. We want to be able to know, analyzing the steady state of this model, what χ, ξ and υ are as functions of the things we know, namely, λ, ρ and θ and other parameters.

Much of what I have to say is like what we did last time so I can go pretty fast.

First of all, from (3a) we know that:

$$\hat{p} = -\sigma\chi.$$

Now let me look at (4a). If I divide both sides of (4a) by p, then I will get another equation for $\hat{p}$:

$$\hat{p} = \rho - \beta K^{\beta-1}(uH)^{1-\beta}H^{\gamma} = -\sigma\chi,$$

from which:

$$\frac{\rho+\sigma\chi}{\beta} = K^{\beta-1}(uH)^{1-\beta}H^{\gamma}.$$

Since ρ is a constant, and σ is a constant (they are parameters of the utility function), and χ in a steady state is a constant, we know that in a steady state the right-hand side is also constant.

Now I want to use equation (1) and to solve it for $\hat{K}$. If I do that I find:

$$\hat{K} = \xi = K^{\beta-1}(\mathrm{uH})^{1-\beta}H^{\gamma} - \frac{Nc}{K} =$$
$$= \frac{\rho+\sigma\chi}{\beta} - \frac{Nc}{K},$$

so (Nc/K) is a constant in steady state, because $\hat{K}$ is. Then:

$$\hat{N} + \hat{c} = \hat{K},$$

and putting that in our notation, we obtain:

$$\lambda + \chi = \xi.$$

Now I want to go back to equation (1). Having found that the right-hand side is a constant, we have the option of differentiating logarithmically and putting that derivative equal to zero. What we get is:

$$(1-\beta)\hat{K} = (1-\beta+\gamma)\hat{H} + (1-\beta)\hat{u}.$$

I already said that I am going to define a steady state like Lucas, as being a situation in which u is constant, so $\hat{u}$ is zero in the above equation. Then we have:

$$(1-\beta)\xi = (1-\beta+\gamma)\upsilon.$$

From this and $\xi = \lambda + \chi$, we get:

$$\upsilon = \frac{(\lambda+\chi)(1-\beta)}{1-\beta+\gamma}.$$

Already we know something about this model that we did not know before. The stock of human capital grows more slowly than the stock of physical capital provided γ is positive, i.e., provided that there is an externality. If there were no externality, and if λ was zero, then the stock of human capital would grow exactly at the same rate as the stock of physical capital. That story tells you that a planner, the social planner maximizing that original integral, will probably have the stock of capital growing as fast as the stock of human capital. The planner would internalize that externality.

I am in the business of solving for all of these growth rates in terms of the parameters of the model. By now you can see that I only need to figure out what χ is. The key is to find out χ. Notice that we are looking for *endogenous growth* rates and also Lucas is looking for *endogenous growth* rates. χ had better be endogenous because every other growth rate just depends on χ and known parameters. If χ were exogenous then this would not be a model of *endogenous growth* after all.

I now have to use something that I have not used yet, which is equation (3c). I am going to take logarithmic derivatives of both sides to obtain:

$$\log(p) + \log(1-\beta) + \beta\log(K) + (1 - \beta + \gamma)\log(H) - \beta\log(u) =$$
$$= \log(q) + \log(\delta) + \log(H),$$

from which:

$$\hat{p} + \beta\hat{K} + (1-\beta+\gamma)\hat{H} = \hat{q} + \hat{H}.$$

Let me find a value for $\hat{q}$:

$$\hat{q} = \hat{p} + \beta\hat{K} + (\gamma - \beta)\hat{H} =$$
$$= -\sigma\chi + \beta(\lambda + \chi) + (\gamma - \beta)\upsilon =$$
$$= \chi(\beta - \sigma) + \lambda\beta + (\gamma - \beta)\upsilon.$$

I also have not used equation (3b). It is very important that now we use (3b) because this is an equation which did not appear in Lucas's original model. It comes from the fact that leisure is a choice variable. Again I am going to do logarithmic differentiation. I find:

$$\hat{N} = \hat{q} + \hat{H},$$

which translates into saying:

$$\hat{q} = \lambda - \upsilon,$$

I am now in the position to solve for χ and you will be surprised. I have three equations involving three unknowns, χ, $\hat{q}$ and υ:

$$\hat{q} = \lambda - \upsilon,$$
$$\hat{q} = \chi(\beta - \sigma) + \lambda\beta + (\gamma - \beta)\upsilon,$$
$$\upsilon = \frac{(\lambda+\chi)(1-\beta)}{1-\beta+\gamma}.$$

Apart from $\hat{q}$, χ, and υ, everything else that appears in the equations is a known constant, a parameter of the model. λ is the rate of population growth, β is a characteristic of the production function, σ is a characteristic of the utility function, and γ is a characteristic of the production function.

You can derive from those equations that:

$$\chi(\beta - \sigma) + \lambda\beta + (\gamma - \beta)\upsilon = \lambda - \upsilon,$$

or:

$$(1 - \beta + \gamma)\upsilon = \lambda(1 - \beta) - \chi(\beta - \sigma),$$

and therefore:

$$(\lambda + \chi)(1 - \beta) = \lambda(1 - \beta) - \chi(\beta - \sigma),$$

i.e.:

$$\chi(1 - \beta) + \chi(\beta - \sigma) = 0,$$

i.e.:

$$\chi(1 - \sigma) = 0.$$

Therefore, in this model, either $\sigma = 1$ or $\chi = 0$. So in general, from this model, we deduce that:

$$\chi = 0,$$

In this case:

$$\hat{c} = \lambda + \mu;$$

although I have ignored exogenous technical progress (i.e. I have set $\mu = 0$) I can put it back here to show that this is exactly the result of the standard neoclassical model. So strangely enough if you alter the Lucas model by allowing a leisure choice, the model reduces to the standard neoclassical model and it provides no *endogenous growth* at all.

I will write down what Lucas's main results are *when leisure does not enter the utility function*. In that case, Lucas find that:

$$\chi = \left(\frac{1-\beta+\gamma}{1-\beta}\right)\left\{\frac{(1-\beta)[\delta-(\rho-\lambda)]}{(1-\beta+\gamma)\sigma-\gamma}\right\}.$$

If we put $a = 0$, so that there is no leisure, and everyone is forced to devote all of his time either to work or to study, then the rate of growth of consumption per head in the Lucas model turns out to be this and he has achieved his goal, that is to say, the rate of growth of consumption per head, the key rate of growth in terms of which all other rates of growth can be expressed, depends among other things on ρ and ξ. So the preference parameters affect the rate of growth and that is what he means by saying that there is an *endogenous growth* there. The preference parameters influence the rate of growth. If, on the other hand, you allow a choice between leisure and work and studying, this turns out not to be so at all and in fact χ becomes zero.

I want to complete the discussion of this model just by adding a few remarks.

The two transversality conditions boil down both to:

$$\rho > \lambda,$$

i.e., the discount rate has to be bigger than the rate of population growth. You can then go back to the model and with some effort work out what $\hat{u}$ and $\hat{l}$ are, that is what is the steady state allocation of time about work and leisure. So the model can be completed perfectly and naturally.

Now let us come back to this rather remarkable difference, that is, to the fact that without a leisure choice χ is equal to what Lucas finds and with a leisure choice χ is equal to μ, i.e., to the exogenously given rate of technological progress.

With a leisure choice, there is an additional margin to be managed in the model and there is not an additional shadow price. The original Lucas's model also has a q and has a shadow price for human capital. But that shadow price is only required to manage one margin. I presume that in the case $\sigma = 1$, which is exactly the case of a logarithmic utility function — the utility function is the logarithm of consumption plus the logarithm of leisure — this problem disappears. The shadow price q that takes care of this equation, that manages properly the margin between study and work, will just also manage the margin between leisure and work. (The logarithmic utility function gives constant "spending" proportions.) But in every other case it will not and the only way that the model can achieve a steady state, in which this equation is satisfied as well as the other equations, so that the three equations that I wrote down here are all satisfied, is by putting $\chi = 0$.

Interestingly enough, what the two young Italian economists have discovered tells me that exactly the same thing is true of Romer's 1986 model. If I can get the time I will discuss that with you. If you look at Romer's 1986 model, at his original model, and insert a leisure choice which is not there in the original model, it too loses the capacity of producing *endogenous growth* rates. (By the way, something similar happens in a very interesting paper by Hahn that is not well known because it is buried in a *Festschrift.*)

Let me go on and raise another problem.

When I described to you what the alternatives were for extending the simple neoclassical model and endogenizing the growth rates, I mentioned three possible ways. One is the accumulation of human capital, and that is what we have just being discussing. There are other models like this but Lucas's model is the father of all of them.

Another way I suggested was that you could give up one of the other assumptions of the neoclassical model. And the one I suggested that is usually given up is the assumption of diminishing returns to capital. I will not say much about that because it has appeared in the literature in rather complicated ways that do not make so obvious what happens.

You remember that all of modern growth theory began with the Domar 1946 model. In the Domar model you come out with the conclusion that the growth of aggregate output is equal to the saving rate divided by the capital-output ratio m:

$$\hat{y} = \frac{s}{m}.$$

This is an *endogenous growth* rate in the sense that Lucas means as endogenous, i.e., the state parameters govern the rate of growth. Anything that govern the saving rate will have something to do with the growth rate. That is true in the original Lucas model: ρ and σ govern the rate of growth.

The production function here is:

$$y = \left(\frac{1}{m}\right)K,$$

because m is the capital-output ratio. This would say that:

$$\hat{y} = \left(\frac{1}{m}\right)\hat{K} = \frac{s}{m}.$$

That is the technology and what we see is that if there are no diminishing returns to capital — and there are not because output is simply proportional to the stock of capital — then the Domar result follows.

There are other ways of getting to that same sort of things. One appears in a paper by Jones and Manuelli (1990), and another in a paper by King and Rebelo (1990). These papers do the same thing in rather different ways.

Jones and Manuelli (1990) operate in the following way. Suppose we measure the capital-labor ratio on the abscissa axis and the output per unit of labor on the ordinate axis. We want to draw a curve which represents a constant returns to scale technology. The usual curve is an increasing concave function. Often in early growth theory one imposes on this function what are called the Inada conditions, namely, that the slope of the function be very large near the origin and very small at the other end. In principle what Inada suggests is that the slope is infinite at the origin and zero at the other end:

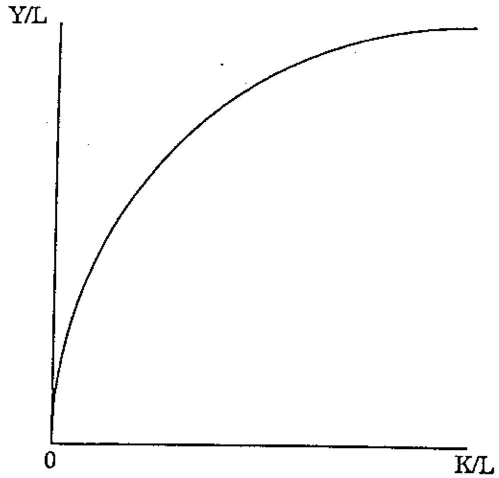

FIGURE 1

Suppose we do not assume the Inada conditions and in particular imagine that the slope of the production function, although always decreasing, has a lower bound. It dimin-

ishes not toward zero, but toward some other number. For instance, we could have a function which is a Cobb-Douglas function plus a constant υ times K:

$$Y = K^{\beta}L^{1-\beta} + \upsilon K,$$

or, in per capita terms:

$$\frac{Y}{L} = \left(\frac{K}{L}\right)^{\beta} + \upsilon\left(\frac{K}{L}\right),$$

with slope:

$$\beta\left(\frac{K}{L}\right)^{\beta-1} + \upsilon.$$

This production function is homogenous of degree one and has positive marginal products. Moreover, it has diminishing marginal productivity because when we come to the second derivatives the added term has already disappeared and so the second derivatives are equal to the second derivatives of the Cobb-Douglas function.

The slope gets smaller as (K/L) goes to infinity. What we have is that this curve is always increasing and concave but asymptotically it goes to a line with slope v. So asymptotically its marginal product is υ. This can behave exactly like Domar's function. If υ is sufficiently large, this curve will have increasing rate of growth everywhere, and the growth model will have an increasing rate of growth everywhere. Asymptotically, as it moves out to the right, it has no diminishing returns. As a result it achieves an *endogenous growth* rate which depends on the saving rate just exactly as in the Domar model. It will not work unless v is sufficiently large. If the non-diminishing returns component of the production function is trivial, then it may not be able to generate *endogenous growth*, but if v is sufficiently large, it will. The simplest way to play this game is not to do the optimization over time, which makes everything to look more complicated that it is, but consider what I call the "behavioristic" version and study the asymptotic behavior.

Now let me give you another example. This example, that is due to King and Rebelo (1990), is a way to get rid of diminishing returns without saying it.

King and Rebelo do something quite different. They suppose that there are two kinds of capital and that the production function for human capital is a function F of the amount of human capital devoted to the production of human capital and the amount of physical capital devoted to the production of human capital:

$$H = F(H_H, K_H).$$

They suppose moreover that the production function for the physical capital is a function G of the amount of human capital devoted to the production of physical capital and the amount of physical capital devoted to the production of physical capital:

$$K = G(H_K, K_K).$$

Each of these accumulable factors of production is produced from stocks of both. Suppose that F and G are both homogenous of degree one, so there are constant returns to scale in both productions, and suppose that the functions are strictly concave. Then, do growth theory with a model like that. For instance, it would be perfectly possible to suppose that some fraction of the output of H and some fraction of the output of K are accumulated:

$$\dot{H} = s_H F(H_H, K_H),$$
$$\dot{K} = s_K G(H_K, K_K).$$

It is easy to show that asymptotically the steady-state growth rate depends on s_H and s_K. What happens in the King and Rebelo model is that there are two capital goods instead of one. But the complex of capital goods producible by constant returns to scale out of the complex of capital goods has by itself no diminishing returns. The essence here is that there is no a primary factor, labor has disappeared.

To conclude, leaving aside the possibility of literally having an endogenous theory of innovations, the other two possibilities, one involving the accumulation of human capital model and the other involving altering the normal assumptions, can both get you *endogenous growth*, but the additional assumptions that you have to make are not trivial and are not obviously true.

In the human capital version it turned out that you have to make two very powerful assumptions. The first assumption was that human capital is produced by human capital and labor in a way that has constant returns to human capital in producing itself. Perhaps I should emphasize that more than I did before.

The Lucas production function for human capital says that:

$$\Delta H = \delta H(1 - u).$$

If you think of this as a technology for producing human capital it has two inputs, human capital and time, and it has diminishing returns in neither of those inputs. It has a constant marginal product of human capital producing human capital and a constant marginal product of time. So the production of human capital does not have diminishing returns in either of these arguments. If you assume that of physical capital we have no problem at all. If you assume that the production function for output is:

$$Y = KL,$$

i.e., if you take a Cobb-Douglas with both exponents equal to one, you would have no problems at all getting *endogenous growth*. But let us suppose that a constant fraction of output is invested, so that:

$$\dot{K} = sKL$$

In this case, it does not take a genius to see that the rate of growth of capital is equal to s times L, i.e., not only it depends on s, but it also depends on the rate at which the labor force has grown. So the first thing that seems to have been done in order to get *endogenous growth* is to make an assumption just about as powerful as that only about the production of human capital. Then, in addition, as we saw today at great length, you have to eliminate the possibility of leisure choice in order to get the main result.

The second point that I just sketch toward the end here, is that there is another way of getting *endogenous growth* without human capital assumption, and that is by finding some way of dropping the assumption of diminishing returns to physical capital, or of diminishing returns to any factor that can be accumulated. Dropping diminishing returns to labor does not help. If you take the inputs that can be accumulated — of course there must be some because otherwise you are not talking about growth at all — and if you arrange it so that in some essential way the group of accumulable factors of production is not faced by diminishing returns — one way to do that is to have no primarily factor at all as in Jones and Manuelli's work — then there can be *endogenous growth*.

THIRD LECTURE

III.1. Introduction

I want to start this hour by going back to where we were at the end of the last lecture before I take up a new topic.

The peculiar result that we found about the Lucas model was that if you allow a leisure choice, that is, if you allow the choice of allocation of time not only between work and accumulation of human capital, but also leisure, then it seems to turn out that the only possible steady state growth rate is the exogenous growth rate, the sum of the growth rate of the labor force and the growth rate of exogenous technological progress. I want to go back over that result and make it clear where that comes from.

First I reproduce the previous lecture three first order conditions for maximizing the current-value Hamiltonian:

(3a) $$c^{-\sigma} = p,$$

(3b) $$Nal^{-\sigma} = q\delta H,$$

(3c) $$p(1 - \beta)K^{\beta} H^{1-\beta} H^{\prime\prime} u^{-\beta} = q\delta H.$$

The first of these conditions just says that the shadow price of consumption has to be equal to the marginal utility of consumption. Then there is another condition that says that when there is a leisure choice the allocation of time must be such as to equalize the marginal value of time devoted to leisure and the marginal value of time devoted to accumulating human capital. q is the shadow price of human capital and what is on the right-hand side of condition (3b) is just the marginal product of time spent accumulating human capital, evaluated at that shadow price. Then, the third first order condition is rather more complicated, but the fact that the right-hand sides of equation (3b) and (3c) are the same tells what the condition means: (3c) equates the marginal value of time spent working to the marginal value of time spent accumulating human capital.

These are the standard results from an intertemporal optimization problem. Anything that can be allocated across a margin has to have a common value at every margin.

From (3a) we deduce that:

$$\hat{p} = -\sigma\hat{c}.$$

(3b) tells us that in the steady state:

$$\hat{q} + \hat{H} = \hat{N},$$

because $\hat{l} = 0$ in steady state by definition.

Then from (3c):

$$\hat{p} + \beta\hat{K} + (1-\beta+\gamma)\hat{H} = \hat{q} + \hat{H}.$$

Finally, there is one more equation that I will write down, a relation that came from doing logarithmic differentiation of the production function with respect to time and making one or two substitutions:

$$(1 - \beta)\hat{K} = (1 - \beta + \gamma)\hat{H}.$$

We deduced from this that:

$$\hat{N} + \hat{c} = \hat{K}$$

i.e., capital per head grows like consumption per head.

That is just to remind you of what happens.

Now, we try to come to the intuitive explanation of why, in the steady state, $\hat{c}$ must be equal to zero except for exceptional cases.

Suppose that $\hat{c}$ is positive, i.e., suppose that consumption per worker is actually growing. Then $\hat{p}$ is negative and therefore $\hat{c} > 0$ implies that p comes to zero: if consumption per head grows, the marginal utility of consumption will go to zero and therefore the shadow price of consumption, namely, the utility price of consumption, will go to zero.

The marginal utility of leisure cannot go to zero because the amount of leisure cannot go to infinity. In the formulation that we have, the amount of leisure goes to a constant and therefore its marginal utility will eventually be a constant. So the only way that the marginal equivalence here can be taking care of leisure is that the adjustment between leisure and human capital accumulation must be taken account of by the shadow price of time. If q has to take on the job of worrying about the marginal equivalence between time spent

in leisure and time spent in accumulating human capital, then there is no shadow price left to take care of the other margin, the allocation of time between work and leisure. That problem does not arise if the rate of growth of consumption is not positive because then the level of consumption is free and p is now available. The mathematics of this says, and this I pointed out fully in the previous lecture, that these five equations are inconsistent unless either $\hat{c}$ is zero or σ is equal to one.

I do not want you to over-interpret this result. I described it to you *as if* it says something disruptive about the Lucas model. That is not necessarily the case. There is after all some path that maximizes the utility integral. It may not approach a steady state (with $\hat{u} = \hat{l} = 0$). We know that if the optimal path approaches the steady state, then either it is the pre-Lucas steady state — the steady state in which $\hat{c}$, $\hat{u}$, and $\hat{l}$ are all equal to zero — or we must be at the rather peculiar situation where σ is equal to one. But, perhaps, there is another path. Perhaps there is a path in which u and l do something else and no steady state is approached at all. The main point of the Lucas model could still be true in the sense that the asymptotic behavior of the optimal path, whatever it is, is endogenous, i.e., it is still influenced by the preference parameters, for instance, which is the result that Lucas was after.

I can indicate to you how that might work if I discuss not an "optimizing" model, but what I have been calling the "behavioristic" model.

Suppose we give up intertemporal optimization. Suppose that what we mean by a steady state in this model is characterized by l and u constant and:

$$\dot{K} = sY,$$

where s is a "behavioral" constant.

Then Lucas, in addition, assumes that:

$$\hat{H} = \hat{h} + \hat{N} = \delta(1 - l - u).$$

This is the key assumption.

The accumulation of capital is a fraction of output:

$$\dot{K} = sK^{\beta}(uH)^{1-\beta}H^{\gamma}.$$

In any exponential path, $\dot{K}$ and K will have the same growth rate, g. The rate of growth of the left-hand side, which will be g, will have to equal the rate of growth of the right-hand side; thus:

$$g = \beta g + (1 - \beta + \gamma)\hat{H},$$

and then that tells us what the value of g is:

$$g = \left(\frac{1-\beta+\gamma}{1-\beta}\right)\hat{H}.$$

This sort of non-optimizing version of the Lucas model tells us that there is a possible exponential rate of growth for K, and therefore also for Y, that must be equal to:

$$\left(\frac{1-\beta+\gamma}{1-\beta}\right)\delta(1 - l - u).$$

What do determine l and u? We are not doing intertemporal-infinite optimization anymore. You can have any theory you like about the allocation of l and u. When you insert it into the above equation you have produced an endogenously determined rate of growth. It is endogenously determined by whatever social mechanism it is that allocates time to work, leisure, and the rest to the accumulation of human capital. We can go a little further if you like. For instance, part of that social mechanism is the investment rate, i.e., the fraction of output that is devoted to ordinary capital accumulation. You could say to yourself that it is possible that the social mechanism that governs s will have something that will be related also to the allocation of time. For instance, suppose goods and leisure are complements. In that case any theory that makes s large, that is the consumption of goods small, is also likely to make leisure small. On the other hand, if leisure and goods are substitutes, if the alternative to consumption is rest, as I hope it is for most of us, then anything that makes s large, that is, the consumption of goods small, is also likely to make l large. Any impulse that makes the consumption of goods small will make the consumption of substitutes for goods larger. So you could in this way construct an endogenous theory of the rate of growth which uses the same technological assumptions as Lucas does but does not try to do what appears to be very difficult to do, i.e., to do this intertemporal optimization of leisure choice. When I say it is difficult to do remember what I mean. I mean that it is difficult to get an interesting steady state if you define a steady state as Lucas defines a steady state. There will be an asymptotic path and it too will depend, in this case, on the taste parameters which are primarily σ and ρ, the intertemporal elasticity of substitution and the rate of time preference. I think you could do just as interesting economics starting in this way. This version of the Lucas story also reveals what is fundamentally the most important point about the Lucas story, and that is that all of the work has been done by this relationship.

Having said that I want to elaborate a little further not on Lucas but on how else you might get *endogenous growth.*

III.2. A General Point about Endogenous Growth Models

I want now to illustrate a general point about all models of *endogenous growth*. I am not going to bother with intertemporal utility maximization but will instead continue with the assumption that a fraction of output is saved and invested:

$$\dot{K} = sY.$$

There has to be an economic and social mechanism that makes that equation true. It is obviously a convenient assumption, but if you want to do economics with it then you have to ask yourself what determines investment and what determines savings in a real situation, and what market mechanism brings them into equality.

I also want to suppose that the production function is of the standard constant returns to scale sort with labor-augmenting technological progress so that we can at least talk about steady states:

$$Y = F(K,AL).$$

Now, I differentiate this with respect to time:

$$\begin{aligned}\dot{Y} &= F_k\dot{K} + F_{AL}(\dot{A}L + A\dot{L}) = \\ &= sF_kY + ALF_{AL}(\hat{A} + \hat{L}),\end{aligned}$$

so that:

$$\hat{Y} \equiv g = sF_k + \frac{ALF_{AL}}{Y}(\mu + \lambda).$$

Now let:

$$\frac{KF_k}{Y} = \beta,$$

so, by constant returns to scale,

$$\frac{ALF_{AL}}{Y} = 1 - \beta,$$

so we have:

$$g = sF_k + (1 - \beta)(\mu + \lambda).$$

Finally, what I can say is that the difference between the rate of growth of output and the exogenous rate of growth $(\mu + \lambda)$ is equal to:

$$g - (\mu + \lambda) = sF_k - \beta(\mu + \lambda).$$

What the literature means by *endogenous growth* is that output should be growing faster than the exogenous factors will make it grow. If output only grows at a rate equal to the sum of the growth rate of the population and the growth rate of the exogenous component of technological change, then this is a model of exogenous growth. There is *endogenous growth* only when the left-hand side, and therefore the right-hand side, is positive. Generally what will make the right-hand side not to be positive is that F_k falls as capital accumulates. Thus, we can say that the job of any model of *endogenous growth* is simply to keep the marginal product of capital from falling too fast as capital accumulates.

There are several possible ways to do this. I want to mention two simple ways in which this might happen.

The first is to suppose that the production function has the special form:

$$F(K,AL) = aK + G(K,AL),$$

where G is homogenous of degree one and behaves exactly like any standard neoclassical production function. In that case:

$$F_k = a + G_k \geq a$$

forever. So if $a > (\mu + \lambda)/s$, there is *endogenous growth*. There are models in the literature which proceed in exactly this way. That is one way of accomplishing the task.

Another more interesting way to proceed, and when I say more interesting I mean that more economics comes out of it, is genuinely to endogenize technological progress. The first paper in that tradition is of course Arrow's paper on *learning by doing*, long ago in the 1960s. What one could do is to suppose that the level of technology depends on the amount of capital that has been accumulated:

$$Y = F(K,A(K)L).$$

That is exactly what Arrow did in the learning by doing paper (1962), where the rate of change of A depends on the rate of investment. We can even imagine the component $A(t)$ to be *external* to the firm, so that accumulation decisions ignore this dependence. (Kaldor's 1957 "technical progress function" was an earlier, and less successful, attempt to accomplish the same thing.)

Now you see what can happen. It is perfectly possible for the quantity $A(K)$ to grow fast enough as capital accumulates to keep the partial derivative of F with respect to the first argument from going to zero.

There are models that work like that and I am going to illustrate this class of models by saying something in this lecture about a paper of Paul Romer called "Endogenous Technological Change" (1990). However, I am not going to go over the model step by step. I want only to show you where the work gets done in this model and I can do that very simply. It is an example of an assumption which performs the function of making A grow rapidly enough with K so as to keep the marginal product of capital from falling too fast as capital accumulation proceeds. So now let me try to show you in a direct and easy way what Romer does.

III.3. Paul Romer's 1990 Model

To lay bare the process of generating *endogenous growth* we can first of all put λ and μ both equal to zero, i.e., we suppose that the population is not growing and that there is no exogenous technological progress. I said that there is *endogenous growth* whenever the growth of output exceeds $(\lambda + \mu)$. So in this model any maintained positive rate of growth of output is *endogenous growth* because λ and μ together contribute nothing. I am also going to suppose, following Romer, that the stock of human capital is constant.

In this model technological progress consists in finding new *varieties* of capital goods, i.e., not so much in making capital goods more productive but in making more kinds of capital goods.

Suppose that at any instant of time there are N varieties of the capital good and that the amounts of the N different types of capital goods that are available for production are:

$$x_1, x_2, \dots\dots, x_N.$$

Although the total amount of human capital available to the economy is going to be assumed constant (H), it is always open to the society to allocate this given stock between the production of output (H_Y) and the production of new varieties of capital (H_A):

$$H = H_Y + H_A.$$

In the steady state of course H_Y and H_A are both constant.

Then, suppose the technology for producing final output has a sort of Cobb-Douglas appearance:

$$Y = H_Y^{\alpha} L^{\beta} \left(\sum_{i=1}^{N} x_i^{1-\alpha-\beta} \right),$$

where L is the constant amount of labor available.

Obviously this has constant returns to scale in all arguments. That looks perfectly routine.

Now, what about manufacturing the capital goods? Romer makes the following assumption which seems to be perfectly reasonable. Suppose that there is some other resource, which we can call R and which might be some special category of labor or some special kind of human capital or something like that, and suppose that it takes η units of this resource to produce one unit of any kind of capital good once it has been invented. It takes ηx_i units of the resource to produce x_i units of the ith capital good and this is true for every i. It would be a cheap generalization to have a separate parameter η_i for each capital good. That would add a little bit of difficulty and achieve nothing new at all. So we might as well take Romer's assumption.

Efficiency and competitive markets will obviously work out so that:

$$x_1 = x_2 = \dots = x_N = \bar{x}.$$

The efficient way to allocate R units of resource to produce capital goods so as to produce output in this way will clearly be to equalize the amount of each of the known capital goods because there are diminishing returns to each of them and only by equalizing the x_i can you equalize the marginal products of the x_i, and since they have the same production technique that is what you clearly want to do. Therefore:

$$N\eta\bar{x} = R.$$

Now let us calculate what the total output will be in this case. Since the product $(H_Y^{\alpha} L^{\beta})$ is constant, let me call it B. Thus, Y will be equal to:

$$Y = BN\bar{x}^{1-\alpha-\beta} = BN\left(\frac{R}{\eta N}\right)^{1-\alpha-\beta} =$$
$$= BR^{1-\alpha-\beta}\eta^{-(1-\alpha-\beta)}N^{\alpha+\beta}.$$

What should strike you is that even with R constant, and H constant, and L constant, it becomes possible for output to be infinitely large as the number of varieties of capital goods goes to infinity. In fact we can go one step further and say that:

$$\hat{Y} = (\alpha + \beta)\hat{N}.$$

Any positive growth rate is *endogenous growth*. Anything that will keep $\hat{N}$ positive, i.e., any economic structure, any market structure, any incentive structure that will keep the number of varieties of capital goods growing, will do the trick.

I feel that I have not emphasized enough how powerful an assumption this technology is. Remember that if N goes to infinity, $\bar{x}$ $(= R/\eta N)$ — the efficient amount of each of those capital goods — goes to zero. Nevertheless, the number of varieties is growing and, as the number of varieties grows, output grows without limits. $\bar{x}$ goes to zero in the course of growth but total output grows. So what that is telling you is that somehow this formulation of the technology makes the productive impact of having a large variety of capital goods very powerful.

I want now to make one purely technical change. I am going to replace sums by integrals. Instead of supposing that there are N discrete varieties of capital goods I suppose that there is a continuum of capital goods running from 0 to A:

$$Y = B\int_0^A x(i)^{1-\alpha-\beta}di,$$

so that, instead of:

$$i = 1, 2, \ldots\ldots, N,$$

we have:

$$0 \leq i \leq A,$$

i.e., we have a density of capital goods.

If you want to maximize the total output subject to the condition that the given amount R is used in the production of different varieties of capital goods:

$$R = \eta\int_0^A x(i)di,$$

then the solution to the problem will say that the marginal product of $x(i)$ must be independent of i and equal to $\bar{x}$. It turns out that:

$$\eta A\bar{x} = R,$$

and this is the exact equivalent of the relationship we got before with A replacing N.

As soon as we have done that, we can replace each $x(i)$ by $\bar{x}$ and then $\bar{x}$ by $(R/\eta A)$ so that:

$$Y = B\int_0^A \left(\frac{R}{\eta A}\right)^{1-\alpha-\beta} di = BR^{1-\alpha-\beta}\eta^{-(1-\alpha-\beta)}A^{\alpha+\beta},$$

and then, with B and R constant:

$$\hat{Y} = (\alpha+\beta)\hat{A}.$$

The next step that Romer undertakes is an exact duplication of Lucas's step. His assumption is that the rate of growth of A is proportional to the amount of human capital allocated to research in discovering new varieties of capital goods:

$$\dot{A} = \delta H_A A.$$

It follows of course that also the rate of growth of output, endogenously determined, is proportional to the *amount* of human capital devoted to research in discovering new varieties of capital goods.

All of the rest of Romer's paper has one function and one function only, namely, to talk about a market structure or an institutional structure that will make H_A constant and positive. Any endogenous mechanism that will keep H_A bigger or equal to some ε which is positive,

$$H_A \geq \varepsilon > 0,$$

will generate *endogenous growth*. Most of the words in the Romer paper are devoted to talking about a complicated structure in which there are firms which manufacture these capital goods, there are other firms which do research on new varieties of capital goods and have a monopoly on the capital goods that they invent and which they then rent or sell to the manufacturing firms. The fact of *endogenous growth* in that paper comes from two things and two things only and you now know what they are.

The point is that the output of the research sector in this economy, which is new varieties of capital goods, is linear both in the human capital input to research and in the number of varieties of capital goods already invented. The key, as Romer says explicitly, is that $\dot{A}$ is linear in A. Suppose that:

$$\dot{A} = \delta H_A A^{\theta}.$$

It turns out that $\theta = 1$ is the only value of θ that will make sense here. If θ is less than one or greater than one something very different happens.

First of all, suppose that:

$$0 < \theta < 1.$$

The rate of growth of A will be:

$$\hat{A} = \delta H_A A^{\theta-1},$$

so that $\hat{Y}$ in this case will be:

$$(\alpha + \beta)\hat{A} = (\alpha + \beta)\delta H_A A^{\theta-1}.$$

What happens as time goes on is that $\hat{Y}$ goes to zero because δ is constant, H_A is bounded by the total amount of human capital, but $(\theta - 1)$ is negative so as time goes to infinity the rate of growth of A goes to zero and therefore the rate of growth of Y goes to zero. So, if θ is less than one asymptotically there is no *endogenous growth*.

What happens if θ is bigger than one? You can surely get *endogenous growth*. Not only can you get *endogenous growth* but if you integrate the differential equation in A for θ bigger than one you will find that A goes to infinity in finite time. There is a time T such that the time series for A goes to infinity and then of course also output will become infinite in a finite time. This outcome does not correspond to common sense. Thus, it is a characteristic of this kind of models that they give the desired result only if θ is equal to one. This is a very special story.

I mentioned to you that the bulk of the paper is to provide an institutional context in which H_A can be understood. When one says that growth is endogenous in the model what that means is that there is a mechanism, an understandable economic market mechanism, which accounts for or is consistent with a positive value for H_A. You could just as well cut through all that if you were prepared to say that the economy has a stock of human capital H that is given and that there is a fraction γ of that that is devoted to H_A. Then the rate of growth of output would just be:

$$(\alpha + \beta)\gamma H,$$

and depending on your interests you could elaborate that a lot.

I do not want you to get the impression that I think that the institutional structure that accounts for H_A is unimportant. It is very important. How quantities like this are determined in a capitalist economy, in a mixed economy, how they might be allocated in a socialist economy, that is what economics is about. But one thing this structure is not about is

endogenous growth. *Endogenous growth* is already taking care of that as long as the mechanism fixes H_A positive.

Now, there are more things that we could say here without going all the way through the intertemporal maximization.

Suppose for instance that W_H is the wage per unit of human capital and P_A is the price at which a new design is sold to a manufacturer of capital goods. We know that W_H will be equal to the value of the marginal product of human capital in producing new designs:

$$W_H = P_A \times \text{Marginal product of } H_A = P_A \delta A.$$

This is a relationship which will hold in a great variety of market structures.

It is also true that human capital can be allocated to the production of goods and so most of the kinds of market structures that we think about will want to make W_H, the rental price of a unit of human capital, equal to the value of its marginal product in the production activity and that will be:

$$W_H = \alpha H_Y^{\alpha-1} L^{\beta} \int_0^A \bar{X}^{1-\alpha-\beta} di = \alpha H_Y^{\alpha-1} L^{\beta} A \bar{X}^{1-\alpha-\beta},$$

where I have used Y as numeraire, so that $P_Y = 1$.

Without any tremendous amount of complication, we have two equations here in W_H and P_A. But we need more equations and Romer's paper adds more relationships, and it gets rather complicated. By the way, I suggest that you might instead just imagine a relationship of the sort $H_A = \gamma H$ and think, in more or less empirical terms, about what would govern the parameter γ. When Romer does that, he gets something a little different from this and provides some insight into the economics of this kind of model. It turns out that the allocation of human capital to the research branch of the economy looks as is shown in FIGURE 2, that is to say, if the total amount of human capital available to the economy is too small, there will be no allocation to research. It turns out in the Romer model, which deduces this from intertemporal utility maximization, that instead of $H_A = \gamma H$, the relationship you get is:

$$H_A = \begin{cases} 0, & \text{if } H \leq H_0 \\ \gamma(H - H_0), & \text{if } H > H_0 \end{cases}$$

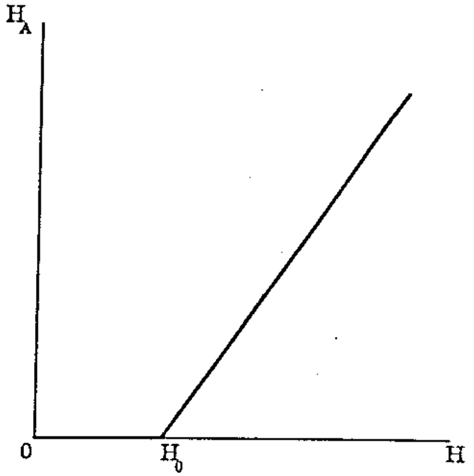

FIGURE 2

You can understand the economics of this. If H is very small, then even if all of the H is devoted to the production side of the economy, the marginal utility of current consumption is so high that the optimizing consumer will choose not to invest in the invention of new kinds of capital goods.

You can of course take the Romer model and do more. For instance, you can add population growth. It would also be possible to write down a wage rate for labor and you could ask yourself what inside this model governs the differential between the price of human capital and price of labor. You could insert several productive sectors and so on. There are many complications that can be built into a model like this, but none of that is *endogenous growth* theory. That is industrial organization, or some other branch of economics. *Endogenous growth theory* consists of this relationship and this particular technology.

What I want to do next, and this I am mostly going to do in the next lecture, is to take this story and turn it on his head and construct the sort of model, analyzed by Grossman and Helpman (1991), in which growth comes about not through increasing varieties of

capital goods but increasing varieties of consumer goods. The industrial structure underlying Romer's model can be simply the industrial structure underlying Grossman's and Helpman's model and so I can use it to show you how the full thing is determined, to write down the complete model which I have not done here.

This kind of model can be neatly represented with the ***Dixit-Stiglitz method*** (1977). It would be fair to say that the Romer technology was simply a transfer into the production side of the Dixit-Stiglitz story on consumption as you will recognize right away.

The idea is the following. Suppose that there are N kinds of consumer goods and suppose that a consumer is consuming the following amounts:

$$x_1, x_2, \ldots\ldots, x_N,$$

of those consumer goods. Then imagine that the consumer has an utility function for consumption which can be represented in the form:

$$U = \left[\sum_{1}^{N} x_N^{\alpha}\right]^{1/\alpha}.$$

Obviously for most utility purposes you do not need the exponent outside but for the specific purposes of Dixit and Stiglitz's arrangement it will prove to be useful to have the utility function homogenous of degree one. Then if we assume that the consumer has a given income and has these goods available at fixed prices, the normal maximization problem gives rise to a very useful outcome. What falls out of the maximization is that if $p_1, p_2, \ldots\ldots, p_N$ are the N prices for these goods, then the nature of the problem is such that a function of these prices can be defined:

$$P(p_1, p_2, \ldots\ldots, p_N),$$

which has the property that P is a true price index, or true index of the cost of living in the sense that the indirect utility, the function that gives the maximum achievable utility, is:

$$\frac{Y}{P(p_1, p_2, \ldots\ldots, p_N)},$$

where, as we will see, the function P has a specific form.

Thus, the above utility function has the property that its indirect utility function can be written in this way and what that means is that the function P of the N prices is a true index of the cost of living. As a consequence, the above ratio is an exact definition of real income. If you know this price index and you deflate nominal income by that price index, you do have real income in precisely the sense that it is the maximized utility that a consumer with this income and facing these prices can actually achieve. You can regard this as a kind of aggregation process.

FOURTH LECTURE

IV.1. The Grossman-Helpman Approach

Instead of pursuing the Romer model — in which growth is achieved through the production of an increasing variety of intermediate goods — I will use some of the same apparatus to illustrate a slightly different approach, due to Grossman and Helpman (1991), in which growth comes about through the combination of two mechanisms: producing an increasing variety of consumer goods with a technology that becomes more productive through the accumulation of knowledge. The "knowledge" is accumulated deliberately. The combination is powerful for reasons already explained.

This model makes use of the Dixit-Stiglitz aggregation procedure. I will only give results, since this is now well known.

We will discuss a model with many consumption goods. We need then an index of consumption:

$$C = \left[\sum_{i=1}^{N} x_i^{\alpha} \right]^{1/\alpha},$$

where, for the purpose of today, I will keep α between 0 and 1. All we need to know at the moment is that the consumption index increases with the consumption of each particular good and it is homogeneous of degree one in all of the xs so that if you double the consumption of each good, you double aggregate consumption. That makes sense.

I want you to consider the following utility function:

$$U = \sum_{i=1}^{N} x_i^{\alpha},$$

where α is between 0 and 1. This utility function is exactly like the consumption index when I treat this as an utility function. As long as α is positive it is just a monotonic transformation. I just want to consider for a moment what happens to a consumer with this utility function who faces given prices and a budget constraint:

$$\sum_{i=1}^{N} p_i x_i = Y.$$

The first order conditions for maximizing the utility are:

$$\alpha x_i^{\alpha-1} = \lambda p_i,$$

where λ is the Lagrange multiplier.

I can eliminate λ right away by noting that these first order conditions say that:

$$\left(\frac{x_i}{x_1}\right)^{\alpha-1} = \frac{p_i}{p_1},$$

and therefore:

$$\frac{x_i}{x_1} = \left(\frac{p_i}{p_1}\right)^{1/(\alpha-1)}.$$

Finally, it is useful if I multiply on the left by (p_i/p_1) so that I have the ratio of the expenditures in x_i and x_1:

$$\frac{p_i}{p_1} \cdot \frac{x_i}{x_1} = \left(\frac{p_i}{p_1}\right)^{\alpha/(\alpha-1)}.$$

Now we have to give effect to the budget constraint. We have:

$$p_i x_i = p_1 x_1 \left(\frac{p_i}{p_1}\right)^{\alpha/(\alpha-1)},$$

from which:

$$\sum_{i=1}^{N} p_i x_i = \sum_{i=1}^{N} p_1 x_1 \left(\frac{p_i}{p_1}\right)^{\alpha/(\alpha-1)} = Y,$$

i.e.:

$$Y = p_1 x_1 \sum_{i=1}^{N} \left(\frac{p_i}{p_1}\right)^{\alpha/(\alpha-1)} =$$
$$= p_1^{1/(\alpha-1)} x_1 \sum_{i=1}^{N} p_i^{\alpha/(\alpha-1)}.$$

I am going now to take the quantity:

$$\left(\sum_{i=1}^{N} p_i^{\alpha/(\alpha-1)}\right)^{(\alpha-1)/\alpha} = P.$$

In this way I have defined a price index. This P is a function of all the prices and it is homogenous of degree one in all the prices. The function on the left, even with $0 < \alpha < 1$, is an increasing function of each price. So this is a price index that falls out of the construction of the problem.

We said that:

$$Y = p_1^{1/(\alpha-1)} x_1 P^{\alpha/(\alpha-1)},$$

Now, at last, I want to solve for x_1. Doing this, I find the demand function for x_1 that corresponds to this utility function. When I do that I find:

$$x_1 = \frac{Y p_1^{-1/(1-\alpha)}}{P^{\alpha/(\alpha-1)}},$$

I want to do one more thing. I want to rearrange the denominator:

$$P^{\alpha/(\alpha-1)} = PP^{-1+\alpha/(\alpha-1)} = PP^{-1/(1-\alpha)},$$

so I can write:

$$x_1 = \frac{Y p_1^{-1/(1-\alpha)}}{PP^{-1/(1-\alpha)}},$$

and now I am at the end. The demand function for x_1, and therefore the demand function for any x_i, is equal to:

$$x_i = \frac{Y}{P}\left(\frac{p_i}{P}\right)^{-1/(1-\alpha)}.$$

The end result is that the demand function for each good takes a very special form. It is real income times a constant elasticity function of the relative price of the ith good where both real income and relative price are defined in terms of the price index P which is the natural, intrinsic, true price index of the problem. So this utility function gives rise to a constant elasticity demand function for each good where the elasticity of demand is the reciprocal of $(1 - \alpha)$, and the relative price of each good and the real income are both defined in terms of the intrinsic price index. I am only going to use a very little bit of this in developing a growth model, but it is useful for you to know it. There is a lot of economics that follows from this neat formulation.

Now I am going to take this for granted. We are going to build a growth model in which there are N consumption goods. The essence of the growth process is going to be, as in Romer, adding to N, i.e., increasing the number of consumer goods that are known in the economy by a deliberate process of research.

I am going to assume that one unit of labor for one unit of time is capable of producing one unit of consumption good for one unit of time. With that, if w is the nominal wage, the marginal cost of producing any consumer good is w. If we imagine that there is a monopoly producer for each of these goods, and we will see how that monopoly comes into existence in a moment, then if the elasticity of demand facing that producer is $1/(1 - \alpha)$, it turns out that the marginal revenue will be equal to α times the price. By the way, this is why we need $0 < \alpha < 1$, so that the elasticity of demand should exceed one. Now we can determine the monopoly price. The marginal revenue, which is (αp), is equal to marginal cost, which is w, so that:

$$p = \frac{w}{\alpha}.$$

The profit of the manufacturer of any one of these commodities is:

$$(p_i - w)x_i = \left(\frac{w}{\alpha} - w\right)x_i = w\left(\frac{1-\alpha}{\alpha}\right)x_i,$$

and total profits, writing X as the sum of all x_i, are:

$$w\left(\frac{1-\alpha}{\alpha}\right)X.$$

Then, the profits of the average firm are:

$$\Pi = w\left(\frac{1-\alpha}{\alpha}\right)\frac{X}{N} =$$
$$= (1 - \alpha)p\frac{X}{N}.$$

Before we use this story to think about growth, there is one more routine thing to be said. The normal way of dealing with a market like this is to add a free entry condition. Suppose that there is a fixed cost of production, and assume that free entry will make N increase until:

$$\frac{(1-\alpha)pX}{N} = \text{fixed cost},$$

i.e., N will adjust so that the profit is zero. This is the usual free entry condition. Here I have put the total cost equal to w times X. I could instead have put the total cost as a fixed cost plus w times X:

$$F + wX.$$

Then the profit for the representative firm would be:

$$\Pi = (1 - \alpha)P\frac{X}{N} - F,$$

and then I could say if this is positive there will be entries. If it is negative, there will be exits and the equilibrium N or the nearest integer is the one for which this is equal to zero.

That is not the natural way to proceed in a growth context. Now I want to convert this into a form which will make a convenient model of growth with the possibility of endogenously determined growth.

The first thing I am going to do is exactly what I did in dealing with Romer's model. I am going always to operate in terms of integrals from 0 to N:

$$C = \left(\int_0^N X(i)^\alpha \, di\right)^{1/\alpha},$$

and that will be the index of consumption that I am interested in, where N, the "number" of varieties, is continuous. Nothing else changes in the story.

The way Grossman and Helpman tell the growth story is that firms change the amount N by engaging in research. There is a technology, so to speak, for producing new commodities. A firm that invents a new commodity holds the monopoly on it forever. So it is entitled to whatever profits can arise. Once the firm holds the monopoly on a variety of the good it can earn its profits for ever. Since a firm monopolizing and producing a new variety of good looks ahead for ever, $v(t)$, the present-value of the stream of profits from t on, is:

$$v(t) = \int_0^{\infty} e^{-rt}\Pi(t)dt.$$

I am not much interested in the equation in this form but I am interested in the fact that the Fisher equation holds at every instant of time:

$$\frac{\Pi}{v} + \frac{\dot{v}}{v} = r, \tag{1}$$

where Π is the profit and r, the interest rate. This says that the own rate of return on the asset v, the profits divided by its current value, plus the capital gain or loss in proportional terms, must be just equal to the interest rate. That is the standard arbitrage equation.

Now we have to go back further and say, how do firms acquire the monopoly power that they hold? How is a new variety of consumer good invented? For that Grossman and Helpman have a technology. What corresponds to free entry is that v, the present-value of the stream of profits from t on for one of this symmetric commodity, must be less than or equal to the cost of creating a new good, which we can call innovation cost:

$$v \leq \text{innovation cost},$$

and v is equal to innovation cost if $\dot{N} > 0$.

What corresponds to free entry, to the zero profits condition, is that the present value of the profits that can be earned from having a monopoly cannot exceed the cost of creating a monopoly, which is the innovation cost, because if it does there will be entries. It could of course be less than the innovation cost but then no new goods would be created. There would be no innovation. As long as there is innovation activity, then the present

value of a monopoly must just be equal to the innovation cost. Schumpeter would of course die at this thought simply because he wanted to play on the fact that there was, at least for an initial interval of time, a real gain, a pure rent to scarce personal entrepreneurship. What is certainly true is that this corresponds exactly to the translation of free entry into this context.

What about the innovation cost? The Grossman-Helpman assumption is that the cost of making an innovation can be described in the following way:

$$\frac{aw}{K_n},$$

where w is the wage, a is a parameter, and where K_n represents the "stock of available knowledge", which is the result of previous research that has entered the public domain.

If K_n were equal to one, then this says in effect that it takes a units of labor to make an innovation, i.e., a units of labor applied to the innovation process will create a marginal extension of the range of known goods.

An alternative way of putting it is that it takes at any instant of time (a/K_n) units of labor to make an innovation. The point is that what Grossman and Helpman intend to do is to treat K_n not as a constant but as something that depends on the accumulated number of innovations that have already been made in the past. It is "learning by doing", but it is an externality. This cost is incurred by the monopolist or by the entrepreneur who then has the monopoly of the new commodity. But if the act of innovation increases K_n and therefore makes research more productive, then there is an additional external effect. For this instant, looking at it microeconomically, i.e., looking at what the incentives for the individual innovating entrepreneur are, we can just say that it is *as if* there are (a/K_n) units of labor which have to be paid.

Next, suppose this economy has L units of labor. I am going to imagine L to be constant so that there is no growth in the labor supply. One could have growth in the labor supply but I am going to cut out all other possible sources of growth but the innovation one, so that any growth is *endogenous growth*.

We can write down an equation for the clearing of the labor market:

$$\frac{a}{K_n}\dot{N} + X = L, \tag{2}$$

where (a/K_n) is the amount of labor it takes to make an innovation; $\dot{N}$, the number of innovations that are currently being made; $(a/K_n)\dot{N}$ is then the total amount of labor involved in the research activity; and X is the total amount of labor involved in the production of goods already known. Since the technology for producing the good is one unit of labor for producing one unit of good, total output is the same as total employment. So if the labor market clears, we have equation (2).

I want to argue next that in order for growth to occur in this model, K_n must be growing through time, i.e., research must be becoming more productive. Suppose, for instance, that K_n were constant. Then equation (2) tells us that $\dot{N}$ and X are both bounded. If $\dot{N}$ is bounded, then $(\dot{N}/N)$ must tend to zero if $\dot{N}$ stays positive. So you could not have a growth rate for N bounded above zero for ever if K_n is not increasing.

I can go one step further. Notice that in this model all of the N commodities known at any instant of time are perfectly symmetric. They have the same technology of production. They enter the utility function in exactly the same way. They have the same elasticity of demand, the same price and even though these are monopolistic competitors, one knows that all of the little x will be equal. So the common output of each commodity will be $\bar{x}$, let us say, and the total output of the commodities will be:

$$X = N\bar{x}.$$

For any given N, we could then write:

$$\bar{x} = \frac{X}{N}.$$

We then evaluate the consumption index C:

$$C = \left(\int_0^N \bar{x}^\alpha di\right)^{1/\alpha} = N^{1/\alpha}\bar{x} = N^{(1/\alpha)-1}X = N^{(1-\alpha)/\alpha}X.$$

That tells you the important fact that the rate of growth of the consumption index is:

$$\hat{C} = \left(\frac{1-\alpha}{\alpha}\right)\hat{N} + \hat{X},$$

so if K_n is constant, and therefore the rate of growth of N must go to zero and the rate of growth of X also for exactly the same reason must go to zero, then the rate of growth of C must go to zero. So if K_n is constant there can be no growth. In fact, if K_n is bounded there can be no growth. As long as equation (2) holds inside this model there cannot be *endogenous growth* unless K_n increases in time. So what we have to assume in order to get *endogenous growth* is that K_n rises though time, that is to say that the innovating activity in addition to producing innovations, which are monopolized by the innovator, also produces an external benefit by making research more productive.

Suppose that K_n is an increasing function of N. Given that the simplest increasing function of N is N itself, Grossman and Helpman, without much further conversation, assume, following Romer, that K_n is equal to N. In this case, from (2), we obtain:

$$a\hat{N} + X = L. \tag{3}$$

If you have listened carefully to what I said in the last lecture, you know that by putting K_n equal to N, a knife-edge assumption has been made. If K_n increases less rapidly than N, then there will not be any *endogenous growth* in this model. If K_n increases more than proportionally than N then there is so much *endogenous growth* in the model that it generates infinite output in finite time. The first part of this sentence makes absolutely sense to me. K_n has to be an increasing function of N in order for there to be *endogenous growth* in the model. That is perfectly true and you can see why. But then to say, oh well, let us choose K_n to be proportional to N, is already saying that nature has been extraordinarily good for the model builder and has made a measure-zero choice of the parameter. This is a story that says that one of the key parameters of this model is just determined so that there can be *endogenous growth*. That is important. But never mind, let us stick with equation (3) of the model.

In the steady state it is going to turn out that:

$$\hat{X} = 0,$$

so that the steady state growth rate of consumption will be just proportional to the steady state growth rate of N:

$$\hat{C} = \left(\frac{1-\alpha}{\alpha}\right)\hat{N}.$$

The next thing that we have to do for this model is to use the Fisher equation. First we have:

$$\Pi = (1-\alpha)\frac{pX}{N};$$

(pX/N) is the total revenue of the firm, the fraction α is its wage cost and the remaining $(1 - \alpha)$ is its profit. The smaller α is, the bigger profits are and that is because the smaller α is, the smaller the elasticity of demand is. We also know that:

$$p = \frac{w}{\alpha}.$$

That comes about because of the one to one technology. At this point it would be simpler if we choose labor for the numeraire and put the wage rate equal to one so that the price is:

$$p = \frac{1}{\alpha}.$$

Here there is of course no loss of generality.

In a steady state in which growth is actually taking place, v, the present-value of the stream of monopoly profits for an innovator, must be exactly equal to the cost of innovation, that is to say:

$$v = \frac{wa}{K_n} = \frac{a}{K_n} = \frac{a}{N}.$$

We are now looking for conditions for steady state and I want a steady state with innovations taking place. I will discuss once we get to the end of the story how there might be a steady state without innovations taking place. At the instant of time at which the number of existing varieties of goods is N, this is the cost of making an innovation and if innovation is actually taking place, then v must be equal to that cost. Under the assumption that K_n is equal to N, it is equal to this formula. The first term of the Fisher equation is then equal to:

$$\frac{\Pi}{v} = (1-\alpha)\frac{pX}{N}\frac{N}{a} = (1-\alpha)\frac{pX}{a} = (1-\alpha)\frac{X}{\alpha a}.$$

What about $(\dot{v}/v)$?

v is equal to (a/N), and a is a constant so that:

$$\hat{v} = -\hat{N}.$$

Nothing mysterious about this. It just says that if innovation is going on all the time, then v must equal the cost of innovation at every time. The cost of innovation is a constant divided by N under the assumptions of this model so the cost of making an innovation is falling like one over N. The cost of innovation is minus the rate of growth of N but the monopoly value of innovation is always equal to the cost and so must be growing at the same rate through time.

Now, I can write the fourth equation of the model which says that:

$$(4) \qquad \left(\frac{1-\alpha}{a\alpha}\right)X = r + \hat{N}.$$

Here I am going to make a simplification that perhaps I should not make but I want to do it anyhow. I am going to go no further with this equation although r itself, the interest rate, is in a way endogenous in a growth model. The underlying parameters would be the parameters of the intertemporal utility function, not this utility function. If I were Romer, or Lucas, or Grossman or Helpman or almost anyone of these people, I would not stop at this point. I would say, well, the interest rate has to be reduced to fundamental taste and technological parameters of the model. The interest rate is a market phenomenon. This is the interest rate at which innovating monopolists discount their future profits, and that should be in principle an endogenous variable. The normal practice in this literature is to suppose that this growing economy behaves *as if* it is acting out the intertemporal utility maximization of the representative consumer with a time preference rate ρ and an intertemporal elasticity of substitution equal to some given constant. Then, we can say what the steady state interest rate must be in relation to those parameters. I find that idea unattractive. I would be prepared to take any theory of the interest rate that anybody has, and insert it here. If you were Böhm-Bawerk, you could put your theory of the interest rate, if you were Wicksell you could put your theory of the interest rate, I do not care. You can think of r here as standing for a function of deep parameters. And if you know what the deep parameters are you put them in there. I will just stop here and not impose as part of a model that ends with *endogenous growth* a particular theory of the determination of the interest rate. What I just said cheats a little bit because among the determinants of the interest rate might very well be $\hat{N}$, the rate of growth of the economy, so there is some simultaneity. But r could be a function of the deep parameters $(N, \hat{N})$ and I would still, in equations (3) and (4), have two equations in X and $\hat{N}$, and I could solve them in principle.

After I have written down the answer in terms of r, I could write down what the answer is in terms of ρ and σ, the deep parameters of the intertemporal utility function. My preference would be just to put everything in terms of the interest rate and invite anyone who is interested in the problem to insert his own theory of the interest rate.

If you let me do that, equations (3) and (4) are two equations in X and $\hat{N}$, and they can be solved. We can look at the following diagram, with X on the ordinate axis and $\hat{N}$ on the abscissa axis, to solve them:

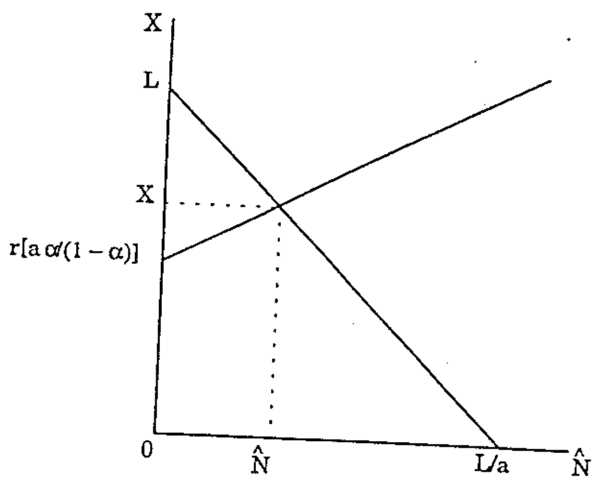

FIGURE 3

Equation (3) is a downward sloping relationship between X and $\hat{N}$ and I will remind you what it means in just a moment. When $\hat{N}$ is equal to zero, X is equal to L and when X is equal to zero, $\hat{N}$ is equal to (L/a). Equation (4) is an upward sloping relationship between X and $\hat{N}$. When $\hat{N}$ is zero, X is equal to $a\alpha r/(1 - \alpha)$. The point where (3) and (4) intersect is the steady state solution.

We also know that:

$$\hat{C} = \left(\frac{1-\alpha}{\alpha}\right)\hat{N}.$$

X is a constant in the steady state here, so $\hat{X}$ is zero, and if you are interested in the rate of growth of consumption you calculate first the rate of innovation — you can think of $\hat{N}$, the percentage rate at which the number of commodities increases through time, as the rate of innovation. The aggregate index of consumption grows at a rate which is proportional to the rate of innovation and the steady state details of the model are solved. Since they are both linear equations, they are certainly trivial to solve. It turns out that:

$$\hat{N} = \left(\frac{1-\alpha}{a}\right)L - \alpha r.$$

Let us think about that for a moment. In this model what makes for a rapid rate of innovation, and therefore for a fast rate of growth of the consumption index is, first of all, the scale. An economy with a large L, i.e., a large economy, will grow faster than a small economy. That is the first thing I want to say about this model. Secondly, anything that makes for a high rate of interest will reduce the rate of growth.

I better come back to the diagram and interpret the slopes of the two curves. Why does equation (3) slope down? Equation (3) slopes down for the simplest possible reason. Labor has to be allocated between research and production. So innovation and production are rival activities. Along this relationship, the more production takes place, the less resources are devoted to innovation and the smaller the rate of innovation is. That is why it is downward sloping. Why is equation (4) upward sloping? There you have to remember where equation (4) comes from. Equation (4) comes from equation (1), from the Fisher arbitrage equation. And if you think it through, the reasoning is this: the faster $\hat{N}$, i.e., the faster the rate of innovation, the more quickly an innovation becomes obsolete, and the faster profits fall. This is a good Schumpeterian point. The more innovational activity is taking place, the faster the monopoly profits evaporate. So, if the rate of innovation is very high, the stream of monopoly profits diminishes sharply. But the present-value of that stream of monopoly profits must equal the cost of innovation. So, if it is falling more rapidly, then it better starts higher. The initial profits must be higher in order that the monopolist can recover the cost of innovation. But we found before that profits were proportional to X. So if the rate of innovation is faster, then the initial level of output has to be higher, so that the initial level of profits can be higher and the present value of profits can be equal to the cost of innovating even though profits are diminishing more rapidly. Therefore along this relationship a bigger $\hat{N}$ goes with a bigger X, with higher initial profits. That makes sense to me. We could then ask why it is that a higher interest rate leads to a smaller $\hat{N}$. Well, a higher interest rate is another way of saying that the present-value of those monopoly profits is being discounted more heavily. That will slow down the rate of innovation by making the innovation less profitable for given X. The fact that the rate of growth is higher when a is smaller is perfectly natural. That just says that the more productive research is — the cost of making an innovation is proportional to a, so if a is small, the cost of making an innovation is very small — you expect there to be more innovation. The fact that a small value of α favors the rate of growth is another Schumpeterian point. That says that a high degree of monopoly favors innovation. The bigger $(1 - \alpha)$, the bigger profits are for any given level of output, so a large value of $(1 - \alpha)$ is a large degree of monopoly in the standard sense that relates to the elasticity of demand. And therefore the bigger $(1 - \alpha)$ is, the faster innovations grow.

Just so that I do not feel that I have not done my duty, I should record the formula for $\hat{N}$ when we suppose that the economy behaves *as if* it is maximizing the integral:

$$\int_0^\infty e^{-\rho t}\left(\frac{C^{1-\sigma}-1}{1-\sigma}\right)dt.$$

If you suppose that the economy behaves so as to maximize this quantity, where ρ is the rate of discount of utility, and $(1/\sigma)$ is the intertemporal elasticity of substitution between goods, if you believe that, then you can relate r to ρ and σ. In that case it turns out that:

$$\hat{N} = \frac{1}{\alpha+\sigma(1-\alpha)}\left[\left(\frac{1-\alpha}{a}\right)L-\alpha\rho\right].$$

The effect of solving for r in terms of ρ and σ is to change the formula for $\hat{N}$ we obtained before into this formula. What is in the square brackets is exactly what we had before with ρ replacing r. But then the whole thing gets multiplied by the factor $[\alpha + \sigma(1 - \alpha)]^{-1}$. You will notice by the way that this formula is exactly right with ρ replacing r if σ is equal to one. σ equal to one is the case of logarithmic utility, so in this case, which is the usual case, this formula works out exactly, where we can think of r as being the rate of time preference. If you like the Ramsey consumer, the intertemporal optimizing consumer approach to this, then a little bit of further work using that standard maximization problem will give you this.

The next step is just to point out that there is another case and that is the case where the previous diagram looks like this:

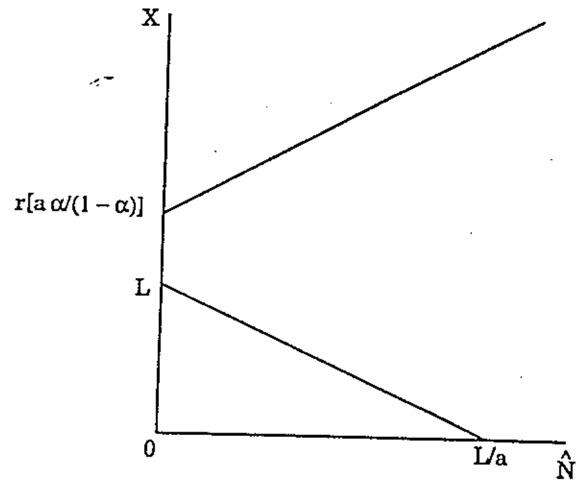

FIGURE 4

The diagram looks like this when it happens that:

$$\left(\frac{1-\alpha}{a}\right)L < \alpha r \text{ (or } < \alpha\rho \text{ in the Ramsey case),}$$

In this case the solution is zero rate of innovation at the point where $X = L$. That is the case where there is no *endogenous growth* in the model at all and the reason why there is no *endogenous growth* in the model is straightforward. The cost of innovation is so high that innovation does not take place. The solution to the model is with all of the resources, the labor or whatever else it is, applied to the production of the existing varieties of goods and no further innovation. That is another Schumpeterian kind of statement.

I think that we have discussed all of the forms of the parameters. The only thing that is added about the comparative dynamics of steady state innovation is that obviously a high rate of time preference will lead to a lower rate of innovation: people, valuing the future very little comparing to the present, will save very little. That is the story that underlies this case. They will save very little and they will be willing to finance less innovation. Summarily, $(1/\sigma)$ is the elasticity of substitution between consumption now and consumption later. If σ is very large, that elasticity of substitution is very small, future consumption is a very bad substitute for present consumption and once again this sort of Ramsey-type economy will be not willing to invest in innovation because innovation pays what is consumption in the future and they are not very happy to exchange present consumption for future consumption.

That is a complete model of a steady state with innovation. I like it because the institutional structure of this economy is simpler than the institutional structure of the Romer model. In the Romer model that we talked about in the previous lecture, the technological progress takes the form of the invention of new kinds of capital goods, or intermediate goods, and there then has to be a complicated market structure.

Finally, I want to re-emphasizes what I said in the previous lecture about this very special kind of model.

The story was that we have:

$$\left(\frac{a}{K_n}\right)\dot{N} = L - X.$$

I pointed out that if we had not casually just said that K_n increases proportionally to N — I took K_n equal to N but it could be bN and just a would be replaced by (a/b), there is no generality there — we would have been faced with a dilemma. I better be explicit about this.

Suppose we have replaced K_n by N^θ:

$$K_n = N^\theta.$$

We know that we want K_n to be increasing with N, but we do not know how fast we want it to be increasing in N, and that is a reasonable flexible function. Let me call B the quantity $(L - X)$ which is constant in the model:

$$\left(\frac{a}{N^\theta}\right)\dot{N} = B,$$

and observe that we now have a differential equation that says that:

$$N^{-\theta}\dot{N} = \frac{B}{a}.$$

If you solve this differential equation, you get:

$$N^{-\theta}dN = \left(\frac{B}{a}\right)dt,$$

i.e.:

$$\frac{N^{1-\theta}}{1-\theta} - \frac{N_0^{1-\theta}}{1-\theta} = \left(\frac{B}{a}\right)t.$$

If $\theta < 1$, then this just says essentially that $N^{1-\theta}$ is exactly a linear function of t and therefore N itself is equal to:

$$N = (C + Dt)^{1/(1-\theta)}.$$

So if $\theta < 1$, a differential equation like this says that N can grow at best at a polynomial rate and therefore its proportional rate of growth is decreasing eventually.

If $\theta > 1$, we would solve in the following way. The simplest think for me to do is to make everything positive:

$$-\frac{N^{1-\theta}}{\theta-1} + \frac{N_0^{1-\theta}}{1-\theta} = \left(\frac{B}{a}\right)t.$$

Then we could say that:

$$\frac{N^{1-\theta}}{\theta-1} = \left(\frac{B}{a}\right)t - \frac{N_0^{1-\theta}}{1-\theta}.$$

$(1 - \theta)$ is a negative exponent, so this is a negative power of N and in fact we would be saying that N to a positive power is equal to the reciprocal of this:

$$N^{\text{positive power}} = \frac{1}{\left(\frac{B}{a}\right)t - \frac{N_0^{1-\theta}}{1-\theta}}.$$

Now you see what I have been telling you. Suppose we start with a rather small value of N_0. When t reaches the value:

$$\left(\frac{a}{B}\right)\left(\frac{N_0^{1-\theta}}{1-\theta}\right),$$

the denominator vanishes. So a positive power of N will become infinite at that finite time and the time path will be ridiculous. Nobody would believe that a growing economy could generate infinite output in a finite time. We would then find out that the only case here where you could create a model of sustained growth is the case $\theta = 1$. This is intended to be an empirical statement. The statement that K_n is equal to N or is proportional to N is seen as an empirical statement, a statement of facts about the nature of innovation. However, it would be very strange indeed if it were so, if that were exactly true. It has to be said that if K_n would be equal to $N^{1.000...1}$, then you could probably live for now to the next thousand of years or so without observing that output was going to become infinite before time came to an end. But nevertheless there is something very suspicious about a class of models that depends on such a very special phenomenon. So there is a research topic here for some of you that are interested in *endogenous growth*. There is a whole class of models of which Romer's is one example and this is another example. Is there a way of formulating them, so that they generate the kind of *endogenous growth* that the authors of these models are looking for, that is more robust? So far I have not yet found such a solution. To tell you the truth, I am not sure that Romer and Grossman and Helpman and other people which do this are aware yet of how special the results are. Maybe they know things that they have not told us. It would be worth to investigate that.

FIFTH LECTURE

V.1. Aghion-Howitt's 1991 Paper

In this lecture I want to describe a paper by Aghion and Howitt (1992). It is technically laborious, so I will not try to cover the details, but you should understand the ideas, because I think this is the general direction in which *New Growth Theory* may have something new to say. In that sense Aghion-Howitt's paper is an example of an interesting tendency.

Their ambition is to make a model that gets close to our intuition about the endogenous generation of new technology. It is still pretty far from anything that feels like real research, academic or industrial. In one way this paper — and the whole literature — may be *too* ambitious. There is probably a substantial exogenous element in the amount and direction of technology change. Fields of research become hot or go dry unexpectedly; in industrial research it is not unusual for results to arise that were not intended when the research was planned and paid for.

Aghion and Howitt introduce several novelties:

(i) They introduce some chance into the R&D process;

(ii) They try to allow for Schumpeter's idea of "*creative destruction*": successful R&D can make the technology invented by previous R&D unprofitable. Thus the rents from successful innovation are temporary. This possibility will be taken into account by entrepreneurs in their decisions about R&D spending. (However Aghion and Howitt do not consider the equally realistic alternative: new R&D can be complementary to previous innovation, and make it *more* profitable at least for a while).

(iii) One of Aghion's and Howitt's results is the possibility of endogenous *cycles* brought about by the innovation mechanism.

Here is a simplified version of their model. There is no capital accumulation, and there is constant employment. There is one final good, produced by labor devoted to final production (x). So final output is equal to:

$$Y = Af(x).$$

It might be better to imagine that labor produces an intermediate good x on a one-to-one basis, and then x produces final output according to $Y = Af(x)$, where $f(\cdot)$ is increasing and concave.

Some labor is devoted to R&D. When successful the innovation is a *new* intermediate good which allows a higher value of A and thus renders the old intermediate good obsolete: no one would use a unit of labor to produce an old x when it could produce a new x instead.

If t refers to the tth innovation (*not* time t), then:

$$\frac{A_{t+1}}{A_t} = \gamma.$$

Suppose n units of labor are assigned to R&D, then innovation arrive according to a Poisson process with arrival rate λn. This means that the probability of an innovation in a given unit of time is equal to λn, the probability of no innovation is equal to $(1 - \lambda n)$, and the probability of two or more innovations is equal to zero. The innovating firm acquires a monopoly on the production of x that is useful until the next innovation. Thus the tth innovation brings a negative externality — it kills the rents of the firm that produced the $(t$-1)st innovation — *and* a positive externality — it makes possible the $(t$+1)st innovation.

Now let V_t stand for the expected discounted rents associated with the tth successful innovation. Let Π_t be the (constant) flow of rent expected by the tth innovator during the profitable life of the innovation, and let ρ be the discount rate for such rents. Then the Fisher equation says that:

$$\rho V_t = \Pi_t - \lambda n_t V_t.$$

The "interest on the value of the innovation" equals the current income — Π_t — plus the expected capital gain — $\lambda n_t(-V_t) + (1 - \lambda n_t)\cdot 0$. (Remember that n_t is devoted to innovation during lifetime of tth innovation, so probability of arrival of $(t$+1)st innovation is λn_t). So we have:

$$V_t = \frac{\Pi_t}{\rho + \lambda n_t}.$$

If there is free entry in R&D, then:

$$w_t n_t = \lambda n_t V_{t+1} + (1 - \lambda n_t)\cdot 0,$$

so:

$$w_t = \lambda V_{t+1}.$$

Notice that a large value of n_t reduces V_t, so Research is like capital investment in this respect: it is discouraged by the prospect of future R&D or investment.

If the constant volume of employment is $\overline{L}$, then the clearing of the labor market means:

$$\overline{L} = n_t + x_t,$$

for every t. Any fluctuations are *not* fluctuations in employment. (This is a major limitation of this model: one of the true risks of R&D is that markets should be weak during the effective life of an innovation, so that it turns out to be unprofitable.)

A successful innovator, monopolizing the intermediate good, faces a demand curve from the final-goods industry:

$$A_t f'(x_t) = P_t.$$

(Use final good as *numeraire*: then the consumption industry demands x until the value of the marginal product — $1\cdot A_t f'(x_t)$ — equals the price of the intermediate good — P_t.)

So the monopolist maximizes:

$$P_t x_t - w_t x_t \text{ (remember the one-to-one technology for producing } x\text{)}$$
$$= A_t f'(x_t) x_t - w_t x_t.$$

It follows that the optimal x_t is a decreasing function of (w_t/A_t), and the best achievable value of (Π_t/A_t) is a decreasing function of (w_t/A_t).

We know that:

$$w_t = \lambda V_{t+1} = \frac{\lambda \Pi_{t+1}}{\rho + \lambda n_{t+1}}.$$

From:

$$\overline{L} = n_t + x_t,$$

we have n_t as a decreasing function of x_t, hence an *increasing* function of (w_t/A_t), say:

$$\frac{w_t}{A_t} = \varphi(n_t).$$

Write:

$$\frac{w_t}{A_t} = \frac{\lambda(\Pi_{t+1}/A_t)}{\rho+\lambda n_{t+1}} = \frac{\lambda\gamma(\Pi_{t+1}/A_{t+1})}{\rho+\lambda n_{t+1}} \text{ (because } A_{t+1} = \gamma A_t.)$$

The left-hand side is equal to $\varphi(n_t)$, an increasing function of n_t. The right-hand side is a decreasing function of n_{t+1} because (Π_{t+1}/A_{t+1}), is a decreasing function of (w_{t+1}/A_{t+1}), hence a decreasing function of n_{t+1}, and the denominator increases with n_{t+1}.

We can write:

$$\varphi(n_t) = \psi(n_{t+1}),$$

with $\psi(\cdot)$ decreasing.

Thus:

$$n_{t+1} = h(n_t),\ h' < 0.$$

In general, as we see in FIGURE 5, there will be a unique steady-state $\bar{n}$ satisfying:

$$\bar{n} = h(\bar{n}).$$

We do not know if n_t tends to $\bar{n}$; it will if $|h'(n)| < 1$ and will converge locally if $|h'(\bar{n})| < 1$. But in any case $\bar{n}$ determines $\lambda\bar{n}$, and also determines $\bar{x} = \overline{L} - \bar{n}$.

In the steady state:

$$Y_{t+1} = A_{t+1}f(\bar{x}),$$

and:

$$Y_t = A_t f(\bar{x}),$$

so:

$$Y_{t+1} = \gamma Y_t.$$

Remember that the index t counts innovations, not time.

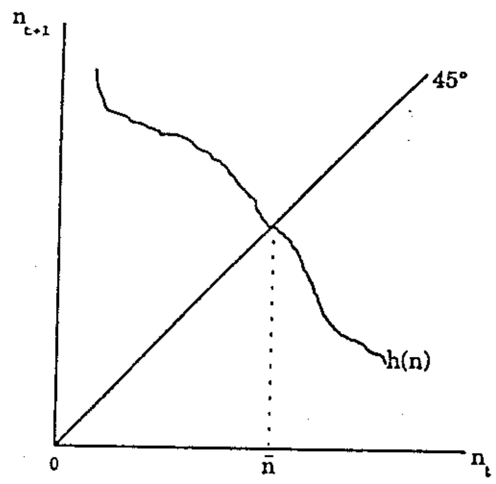

FIGURE 5

Let $Y(z)$ be output at *time* z. Then:

$$\frac{Y(z+1)}{Y(z)} = \gamma^{\varepsilon(z)},$$

where $\varepsilon(z)$ is the *number of innovations* occurring in the steady state between time z+1 and time z. So:

$$\ln Y(z+1) - \ln Y(z) = \text{the rate of growth of output in steady state} = \varepsilon(z)\ln\gamma,$$

where $\varepsilon(z)$ has a Poisson distribution with mean $\lambda\bar{n}$. Thus the rate of growth of output in steady state (E) is:

$$E = \lambda\bar{n}\ln\gamma.$$

One can say that the logarithm of output in steady state follows a random walk with drift $\lambda\bar{n}\ln\gamma$. Thus the *expected rate of growth* is proportional to $\bar{n}$. This is an endogenous rate of growth and it will depend on anything that helps to determine $\bar{n}$. Any tax incentive or regulation or subsidy that increases the level of resources devoted to R&D will increase the expected rate of growth (and also the variance of the growth rate, equal to $(\ln\gamma)^2\lambda\bar{n}$ by the Poisson property).

There are some important comments to be made about this model.

(a) Notice the implication that the *growth rate* is increasing in $\bar{L}$, the scale of the economy, because then $\bar{n}$ will be larger for any tax-subsidy situation. Usually when we compare the R&D intensity of different economies, like Japan, US, EC, we look at R&D *spending*. This has always seemed foolish to me. A dollar of R&D should produce the same amount of innovations in a small economy as in a large one. The model confirms that. If the US has the same R&D-GNP ratio as Japan, it should do better than Japan because it is larger. Of course, realistically, innovations diffuse internationally, so much of this is irrelevant.

(b) There is a large element of arbitrariness in the model, which may be hidden because the model is interesting. The endogenousness of the growth *rate* is in a sense merely assumed. Each innovation makes:

$$\frac{Y_{t+1}}{Y_t} = \gamma > 1.$$

If instead each innovation increased A_t to:

$$A_{t+1} = A_t + \gamma,$$

then one would have something like:

$$Y(z+1) = Y(z) + \lambda\bar{n}\gamma,$$

and thus the rate of growth,

$$\frac{Y(z+1) - Y(z)}{Y(z)} = \frac{\lambda\bar{n}\gamma}{Y(z)},$$

would tend to zero as Y(z) tends to infinitive. There is no sustainable purely endogenous growth rate.

(c) An exogenous element could be introduced by making λ (the productivity of R&D effort) a function of calendar time, or of something else that measures the difficulty of making a productive innovation. Or else γ, the "size" of an innovation, could be non-constant. A more radical change would be to abandon the Poisson assumption, which says that the probability of making an innovation of given size depends only on n, independent of past history of innovation. This is not unreasonable: one can think of reasons why past research should make success less likely (easier innovations picked off first) and also reasons why past research should make success more likely (accumulation of basic science). There are "breakthroughs" that open up a whole new field for innovation, gradually exhausted until next breakthrough occurs. This is harder to model. In the meanwhile, Aghion's and Howitt's is a real step forward.

(d) If n_t is itself cyclical, then fluctuations can have permanent effects. A temporary increase in n will raise productivity. This will not be forgotten when n diminishes so productivity remains permanently higher because of a one-time innovation.

Will regular, repeated fluctuations in n increase or decrease the *average* rate of growth? That depends on whether arrival rate of innovations is convex or concave in n. The λn case makes fluctuations neutral.

There are alternative assumptions about n as well. It is possible that n might increase in recession, because fewer resources are needed for production, and because competition is more intensive. This is outside the formal model. Some indeterminacy is probably desirable, to leave room for "animal spirits". Of course expected recession reduces the profits from innovation and would probably reduce n.

(e) The basic difference equation,

$$n_{t+1} = h(n_t),$$

even with $h'(n) < 0$, can have a variety of solutions. A more complex model with h' changing sign could have several steady states and even chaotic solutions. One special case occurs if $h'(n) = -1$ for an interval of n around $\bar{n}$. That gives a regular cycle of period two (like the cobweb).

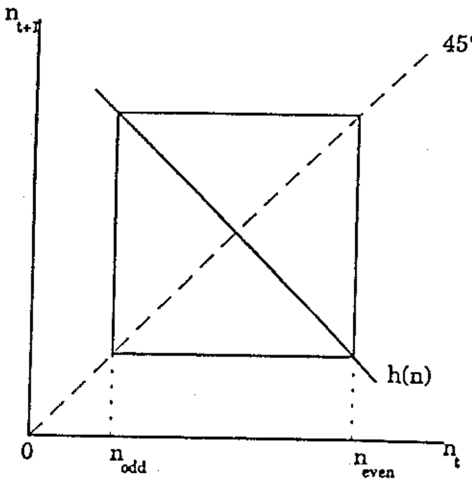

FIGURE 6

When n is low, innovation is expected to be profitable, so n is high next period; but that gives rise to expectation that innovation will be unprofitable, so n is low in the following period. A steady state with $n = 0$, and zero growth, is possible, even under rational expectations: it is correctly anticipated that positive n_{t+1} will reduce Π_{t+1} enough so that the R &D expenditure $(w_{t+1}n_{t+1})$ will not pay off.

This is a promising line of thought, but what is needed are empirically-based hypotheses about the stochastic process that describes innovation. Outcomes can be simulated.

SIXTH LECTURE

VI.1. Endogenous Growth and Empirical Work: An Introduction

I want now to introduce a bit of discussion of empirical work. I do that for *two* reasons.

First of all, because I believe very strongly that work on growth theory ought to be integrated with empirical work.

Secondly, I also thought it was worthwhile because by a coincidence what has happened is that a particular sort of econometric study has become very popular along with the study of models of *endogenous growth*. What happened was a very interesting kind of coincidence which is that along about the time that Lucas (1988) and Paul Romer (1986) were reviving the notion of *endogenous growth*, there came into existence a body of data which had not previously been available and which asks to be made the basis of empirical analyses of growth. These are the Summers-Heston data.

What Summers and Heston (1984, 1988, 1991) have done is to compile the main variables from the National Income Accounts of about 120 different countries, all of the members of the United Nations so to speak, for the period from 1960 to 1985 and to deflate these data to 1985 on a comparable basis using as close to purchasing-power-parity prices as they can. They are available from 1960 to 1985, perhaps not for every year in that period but for benchmark years in that period: real GNP, real investment, real GDP, real consumption expenditure, all of the key variables from the National Income Accounts on a comparable basis, deflated in an intelligent way not using the official exchange rates or anything like that, but making at least an attempt to achieve purchasing-power-parity prices. Since these data are available from 1960 to 1985, it is possible to calculate for each of these countries the average rate of growth of real GDP over a quarter century period, and it is just irresistible to use these data in regression equations in which g_i is the average rate of growth of real GDP for country i during the period from 1960 to 1985, expressed as a function of variables that might be associated with the growth rate. If you accept the idea of endogenously determined growth rates, then the Summers-Heston data give you the chance of testing empirically what are the characteristics of each country that are associated with the growth rate. If growth rates are endogenous, you can test hypotheses about causal variables or about potential causal variables on the right-hand side and in fact over the past seven or eight years very many such cross-country regressions have been done with the growth rate as the left-hand side variable, as the variable to be explained in the regression, and all sorts of things on the right-hand side. I will discuss that a little bit. I am only going to talk about a couple of the papers that have been written along these lines in any detail. These two studies, which come out on different sides of the question, strike me as very interesting. They illustrate what seems to me to be the main fact about these empirical studies, mainly that *they are not robust*. Quite different econometric answers to important questions can be obtained from quite minor changes in the econometric approach: slightly different periods, slightly different choices of the countries to study, slightly different models, slightly different functional forms, and so on. Unfortunately, no very powerful empirical generalization emerges from this.

Apart from the general question of causality, the kind of questions that can be asked of these cross-section studies, and that have been asked, are like the following: is the growth rate associated with the ratio of investment to output in country *i*?

$$g_i = F\left(\frac{I_i}{Y_i}, \ \ldots\ldots\right).$$

You see why that is an interesting question. In the simple neoclassical model of growth the answer is 'no'. Asymptotically, in steady state, the growth rate is not associated with the fraction of output invested, only the level of output matters. In an *endogenous growth* model, and I will give you an example of this, the answer would be 'yes'. Even if we take account of other variables, there should be a significant partial correlation of the growth rate with the investment quote. That is one question to ask with the idea of distinguishing between endogenous and exogenous theories of growth.

Another question that is often asked is the following: is there convergence of growth rates? Or, is there evidence that different countries, although they have different growth rates now, perhaps are moving toward a common growth rate? Or, is there evidence that different countries, although they have very widely different income levels now, are developing in such a way that you would predict that income levels will converge? The way this is usually done in the empirical literature is that in a cross-country regression like the following, in which the growth rate is regressed against a group of right-hand side variables, one of the variables might be the level of output at the beginning of the period (Y_i^0):

$$g_i = F\left(\frac{I_i}{Y_i}, \ \ldots, \frac{Y_i^0}{L_i}, \ \ldots\right).$$

It is taken as evidence of convergence in some form if the coefficient in the regression of g_i on the variable (Y_i^0/L_i) is negative. That is to say, if it turns out that after taking account of other variables, it remains the case that poorer countries grow faster than richer countries, then that is taken as an evidence of convergence. However, this is by itself a very weak evidence of convergence. For instance, an alternative hypothesis is that there are several sub-groups of countries, and that there is convergence (either of growth rates or levels of income) within each group but not convergence between groups. In that case even if you find a predominant negative coefficient, in a regression like that, it might simply mean that within each group, poor countries do grow faster than richer country, but it is missing the fact that between groups there may be no association of that kind so there is no tendency for the groups to converge or to catch up with each other.

Another important question which emerges is the following. In *endogenous growth* models, the normal sorts of policy that we think about can affect the growth rate, whereas in the neoclassical model the normal kinds of policies, tax policies of one sort or another, will have an effect only on the level of income and not on the growth rate. The obvious example of what I mean is the Lucas model where the growth rate depends on the amount of time that is not devoted to production and not devoted to leisure, but devoted to the accumulation of human capital. If that determines the growth rate, as it does in the Lucas model and does in the Romer model as well, then it is certainly within bounds that tax policy could affect the fraction of time that is devoted to the accumulation of human capital. Any subsidization of whatever activities we think of as leading to the accumulation of human capital will certainly encourage more time being spent on the accumulation of human capital. In the Lucas model that affects the growth rate.

There are models in which physical capital accumulation affects the growth rate unlike the neoclassical model and we certainly all believe that governments can subsidize investments and increase the fraction of output invested. It is a very important difference between the two classes of models that in the family of *endogenous growth* models standard kinds of policies can affect the growth rate. So another typical thing that happens in this regression analysis is that policy variables are inserted in the right-hand side to see whether the cross-sectional regression can detect any dependence of the growth rate of the country on its policy variables, for instance, on the level of government consumption relative to GNP, or on the tax rate on investment. Different authors of papers insert different policy measures of tax policy, or subsidy policy on the right-hand side of cross-country regression to see if they can detect a permanent effect of those things on growth rates:

$$g_i = F\left(\frac{I_i}{Y_i}, \dots, \frac{Y_i^0}{L_i}, \dots, \text{Policy Variables}\right).$$

Barro (1990) even includes variables which are intended to measure political stability — the number of political assassinations, or things like that — on the right-hand side to see whether there is a permanent association between political stability and growth.

All sorts of things have been done. However, as I said, they tend not to give robust results.

With that introduction, I want to talk about two papers just to describe to you what they sound like and what sort of interesting results you can get. I have chosen two papers which come up with rather different conclusions just because I want to illustrate to you the fact that a series of scholars working on this can in fact come up with quite different conclusions.

VI.2. De Long-Summers's 1991 Paper

The first paper I want to describe is by De Long and Summers. It appeared in the *Quarterly Journal of Economics* in May 1991. I shall also tell you that that issue of the Journal also contains an article by Summers and Heston which gives a lot of their data such as very convenient tables of GDP per worker in real terms, rates of growth, and things like that.

What De Long and Summers do is the following. The left-hand side variable in their cross-country regressions is g_i, the real growth rate of country *i*. They work with two samples of countries. One sample which consists of twenty-five relatively rich countries, where relatively rich is defined as GDP per worker greater than 1/4 of GDP per worker in the US. In the Heston-Summers sample there are twenty-five countries that have a GDP per worker which is larger than 25% of the GDP per worker in the US and that is one sample with which De Long and Summers work. The other sample consists of sixty-one countries, where the other thirty-six countries are poorer than the first twenty-five countries.

What De Long and Summers want to demonstrate, and think that they do demonstrate, is that there is a very important *causal* relationship between the growth rate of country *i* and investment in machinery and equipment relative to GDP. In other words, this is one of those papers which is trying to establish an association which they treat as causal.

They include in the regression, after having tried some other things, four variables simultaneously. One is investment in machinery and equipment. Another I will just describe as "*gap*", and that is the percentage by which the US GDP per worker at the beginning of the period exceeds the GDP per worker in the particular country. This is intended to test the hypothesis that poor countries grow faster than rich countries. The other two variables that are included are labor force growth and other investment divided by GDP — physical investment, not human capital investment in any case. In effect, the basic statistical exercise for the sample of 25 countries and for the more inclusive sample of 61 countries is to run a linear regression in which the growth rate for a country is regressed against the intensity of investment in machinery and equipment, the gap between the given country level of productivity and the US, the growth of the labor force, and the intensity of investment in other forms of physical capital. What De Long and Summers find is that it is only the first two variables that have statistically significant regression coefficients. The other two variables have statistically insignificant regression coefficients. They do not enter the regression analysis strongly. The coefficient on the gap is positive as they formulate it and that means that poor countries grow faster than richer countries, certain other things equal. They do find that investment in machinery and equipment enters with a strongly significant positive coefficient, whereas other forms of investment do not enter in a statistically significant way at all. That is the basic empirical result of De Long-Summers's paper. When this is extended to the larger sample of 61 countries, the relation becomes considerably weaker. Among the 25 countries the R^2 is 0.66, whereas for the 61 countries the R^2 is 0.29. But the general shape of the relationship remains the same. When it is extended to the larger sample there is more residual variance obviously but it remains true that the only two significant variables are investment in machinery and equipment and the gap.

The economic importance of the coefficient of the investment in machinery and equipment can be described in the following way. The regression equation says that a country that increases the fraction of its GDP that is invested in machinery and equipment by 3 percentage points will add one percent per year to its growth rate. If you look at the frequency distribution of these 25 countries according to the fraction of GDP invested in machinery and equipment, that frequency distribution has a standard deviation of 0.03. So the country that moves from being one standard deviation below the mean to the mean, by changing the fraction of GDP invested in machinery and equipment from 0.06 to 0.09, or something like that, would add 1.02 percent for a year to its growth rate. That is very large. Over a 25-year period, that adds 29% to the GDP. The calculation that De Long and

Summers make and which they believe — I am skeptical about all of this — implies that the social rate of return to investment in machinery and equipment is of the order of magnitude of 30% per year. There must then be a very strong external effect because the return on investment in machinery and equipment to private businesses is not even near to 30%. If therefore De Long and Summers are to believe this, they must believe that investment in machinery and equipment brings with it a positive externality that other forms of investment do not. And what they suggest is that the externality takes the form that technological innovation is transferred through machinery and equipment. Investment in buildings, for instance, provides buildings but it does not for the investing country transfer technology from wherever the technological frontier is, presumably in the US or in the advanced countries of Europe or Japan. On the contrary, investment in machinery does.

Now I just want to say that that by itself is not a foolish idea. I have talked to a lot of business people and manufacturers in the US and asked this question and they said that it makes sense to them, that often new technology is transferred into a company by the suppliers of machinery and equipment. Obviously, they would not effectively apply that technology if they were not investing in machinery and equipment. So it is not a senseless idea. However, whether it can account for the difference between the 10-15% private rate of return and the 30% social rate of return to investment in that kind of capital seems to me to be doubtful.

If you accept De Long-Summers's regression results, at least *two* important conclusions follow.

First of all, it is certainly strong evidence in favor of *endogenous growth*. If you can affect the growth rate by affecting the fraction of GDP invested in machinery and equipment, then that is certainly *endogenous growth*. As against that it has to be said that what we are looking at are growth rates averaged over a single 25-year period and this is probably not long enough to reject the hypothesis that a maintained increase in that ratio would lead initially to an accelerated growth of 1% per year, but then in the next 25 years the increment of the growth might afterward fall to a third 1% per year. So there still the possibility that over a period longer than 25 years the increment to the growth rate could not be maintained. So it is not an absolutely convincing confirmation of theories of *endogenous growth* but it certainly goes in that direction.

The *second* implication from this result is certainly that the tax system and regulatory system should favor investment in machinery and equipment far beyond what is done.

The next thing that has to be discussed in connection with this paper is the direction in which the causality goes. De Long and Summers are perfectly well aware that one's natural inclination is to say that maybe the causality runs the other way, that is, maybe it is

not that a high ratio of investment in machinery and equipment to GDP increases the growth rate. It could be that the countries that are growing rapidly for some other reason invest a lot in machinery and equipment. They make several intelligent arguments against that interpretation and in favor of their own preferred interpretation, which is that causality runs from investment in machinery and equipment to high rate of growth. The most important piece of evidence that they offer in favor of this direction of causality is really extraordinarily interesting. They find that the countries in the sample differ quite a lot in the relative price of machinery and equipment. If you look, for each of these countries, at the price level for machinery and equipment deflated by the GDP price level, there is quite a lot of variation. In some countries in the sample, machinery and equipment are relatively expensive. In other countries in the sample, machinery and equipment are relatively cheap. Suppose some countries grow rapidly and *therefore* they invest a lot in machinery and equipment. They would say that in those countries with high growth rates the price for machinery and equipment ought to be relatively high. There would be a strong demand for investment in machinery and equipment and that should push the price of machinery and equipment high, whereas, they say, if the causality runs the other way, it is more likely that the relationship between growth and prices of machinery and equipment would be different. What they find is that very strongly the rapidly growing countries are countries which have a low level of price of machinery. So it appears to be a supply-side phenomenon rather than a demand-side phenomenon. In fact, apparently, you could use the relative price of machinery and equipment in the basic regression instead of the fraction of GDP invested in machinery and equipment and the regression coefficient will be negative of course rather than positive, but the fit will be almost as good. They argue that in a very strong way, the message that has been sent by their data is that what is good for growth is a low machinery and equipment price such that there is a lot of investment in machinery and equipment and therefore a lot of growth. That would seem to suggest that a good policy tool for taking advantage of this relationship would be simply a tax credit for investment in machinery and equipment.

They make one other argument for their direction of causality. They suggest that, if the causality went from fast growth to high investment in machinery and equipment, then it should not matter what the source of growth was and for instance rapid labor-force growth should be positively associated with high investment in machinery and equipment. If fast growing countries invest a lot, then countries which are growing fast just because the population is growing fast should also invest a lot. That turns out not to be so, that is, there is no correlation between the rate of labor force growth and the rate of investment in machinery and equipment. So they conclude that the causality is the way that an *endogenous growth* model would suggest.

What are the main statistical or economic weaknesses in this argument? There is one statistical weakness that they point out themselves. If in the regression with the 25 relatively rich countries you add a dummy variable for the continent in which the country is — there are only four, North-America, Europe, Asia and Latin America, because none of the African countries fit into the 25 rich countries — then what happens statistically is very interesting and I want to describe it to you. In the original regression for the 25 rich countries without the dummy variables for the continents, the coefficient of the investment ratio is 0.337 with a standard error of 0.05. When you put in the dummy variables for the four continents, this coefficient falls to 0.053 (with a standard error of 0.063), that is, it is no longer statistically significant and what we find is that there is a large positive coefficient of the dummy variable for Asia, a small positive coefficient of the dummy variable for Europe, and a negative coefficient on the dummy variable for Latin America. From that point of view, you can argue that what this regression is telling you is that Japan grows faster than Argentina, that being in Asia is better than being in Latin America, and if you take account of that the rate of investment in machinery and equipment plays no role. However, in the larger sample where the regression works less well in general, if you put in the five (because now there African countries) continental dummy variables here, they do nothing at all. They leave the coefficient of investment in machinery and equipment unchanged, and the dummy variables are not statistically significant themselves. De Long and Summers interpret this as saying that the reason why in the smaller sample of countries the continental dummies outperform the investment ratio is that this sample of rich countries is so small that within each continent there is not enough variation in the investment ratio to teach the regression anything. But if you enlarge the regression to the 61 countries, so you have some countries that are growing rapidly, some countries that are growing very poorly, and you can observe that within the Latin America the slow growing countries have a low ratio of investment in machinery and equipment and the fast growing countries have a high ratio of investment in machinery and equipment, then the regression equation is able to distinguish that. That is plausible. On the other hand, it has been found that the large-sample results are strongly influenced by the data for one African country, Botswana, with very rapid growth and very high investment (in the diamond industry). This is very suspicious.

There is a second general statistical point that I can make about this but, as you will see, it is incomplete. A graduate student in my University, at MIT, by the name of Charles Jones (1992), is interested in this sort of things as well. He has found that for the main countries, the countries for which you can do the analysis, the series of annual growth rate $g_i(t)$ is a stationary series, a trendless series, for the period of time from 1960 to 1985. There is no trend, neither deterministic nor a unit-root trend, in the growth rates. Therefore, he reasons, any other time series that is to have a permanent effect on the growth rate ought to be trendless as well. The ratio of total investment to GDP in these countries is not trendless and so Jones's suggestion was (he did this work before he had read De Long-Summers's paper) that you cannot make a case that total investment ratio to GDP affects the growth rate because it is not stationary and the growth rate is stationary. Jones is now trying to test the stationarity of the De Long-Summers time series but that is harder to put together in an annual basis and it will take some time. But, I suspect, and he suspects, that he will find that this series is not trendless in these countries and if it is not, then there is a difficulty in the De Long-Summers statistical procedures.

I want to say only one more sentence about this and that is that I have some skepticism about the De Long-Summers result because it is so strong. If they were to tell me that the social rate of return in equipment investment instead of being 30% per year was 15% per year or 16% per year I could then have quite agreed. But that it is 30% per year I would require more convincing reasons because I would have been asked to believe a very powerful assumption.

Now let me drop this and move on, and discuss another paper which produces evidence almost in exactly the opposite direction, that is, against the *endogenous growth* model.

VI.3. Mankiw-Romer-Weil's 1991 Paper

This is a paper by Mankiw, Romer — not Paul, but David, a different Romer — and Weil and is called "A Contribution to the Empirics of Economic Growth" (1992).

The paper sets out to see how well the Summers-Heston data can be explained by a straightforward neoclassical growth model without any endogenousness. They conclude that they can, with one modification which still leaves the model a purely *exogenous growth* model. You see just how inconclusive this whole body of empirical work can be.

Let me remind you a little bit about neoclassical growth theory. Suppose we take a model which says that output is a function of the capital stock and of the input of labor with labor-augmenting technological progress:

$$Y = F(K,AL),$$

and suppose that we add that investment is a fraction of output minus depreciation:

$$\dot{K} = sY - dK,$$

where d is the rate of depreciation. $\dot{K}$ is net investment, and $\dot{K}$ plus dK is gross investment. Then:

$$\dot{K} = sF(K,AL) - dK,$$

and we use the constant returns to scale property to say that:

$$\dot{K} = sALF(z,1) - dK,$$

where $(K/AL) = z$, which is capital per worker measured in efficiency units.

It follows from this that the growth rate of the capital stock is:

$$\frac{\dot{K}}{K} = \hat{K} = s\frac{1}{z}F(z,1) - d,$$

and if we now adopt the Cobb-Douglas form, z^α, $\hat{z}$, the rate of growth of z, is:

$$\hat{z} = \hat{K} - \hat{A} - \hat{L},$$

where:

$$A = A_0 e^{gt},$$

and:

$$L = L_0 e^{nt},$$

so that:

$$\hat{z} = \left(\frac{s}{z}\right) z^\alpha - (d + g + n),$$

from which:

$$\dot{z} = sz^\alpha - (d + g + n)z,$$

The next step is to ask what is the steady state value of z, and that is easily calculated putting $\dot{z}$ equal to zero:

$$z^* = \left(\frac{s}{d+g+n}\right)^{1/(1-\alpha)}.$$

Now what I want to do is to put that back in the production function and to see what the steady state value for Y is. I am doing all this because I am deriving in this way the regression equation that Mankiw, Romer, and Weil actually fit. In steady state:

$$\frac{Y}{L} = AF(z^*,1) = Az^{*\alpha} = A\left(\frac{s}{d+g+n}\right)^{\alpha/(1-\alpha)}.$$

Let us take logarithms:

$$\log\left(\frac{Y}{L}\right) = \log A(0) + gt + \left(\frac{\alpha}{1-\alpha}\right)\log s - \left(\frac{\alpha}{1-\alpha}\right)\log(d+g+n).$$

This is the regression that Mankiw, Romer and Weil fit across countries. The left-hand side of the regression is the average, over the all period, of the logarithm of output per worker country by country. $\log A(0) + gt$ is a constant because the value of t is the same for each of these observations, the observations of different countries. The assumption here is that d and g, the growth rate and the depreciation rate, are the same for each country but the population growth rate differs for each country so that the value n has a subscript i. Then the regression equation is:

$$\log\left(\frac{Y}{L}\right)_i = \bar{A} + \left(\frac{\alpha}{1-\alpha}\right)\log s_i - \left(\frac{\alpha}{1-\alpha}\right)\log(d+g+n_i) + \varepsilon_i.$$

The model, which is the simplest model you can imagine, has at least one prediction. It predicts that the coefficient of the variable ($\log s_i$) will be the negative of the coefficient of the variable $\log(d+g+n_i)$. When the three authors carry out that regression, they find first of all that it is a fairly good fit, the $\bar{R}^2$ is about 0.6. Secondly, the standard test of the hypothesis that the coefficient of the variable ($\log s_i$) is equal to minus the coefficient of the variable $\log(d+g+n_i)$ is passed. However, when this regression is re-estimated imposing the hypothesis that the two coefficients are the same except for the sign, the estimated value of the coefficient is 1.43 which leads to an α of about 0.6. That goes against the model because we would expect that α, which is the capital elasticity in a Cobb-Douglas production function, is of the order of magnitude of 0.3. So as an intermediate step in their work what Mankiw, Romer and Weil find is that 'yes', the regression passes the test of the two coefficients to be equal but it gives an unsatisfactory value for α. They then go on to do something else. They say, why do we not introduce human capital since that plays such a big role in *endogenous growth* models? Why do we not introduce human capital into this model simply by enlarging the production function? Why do not we ask how the data would react to a model which says that:

$$Y = K^{\alpha}\, H\, (AL)^{1-\alpha-\beta},$$
$$\dot{K} = s_K Y - dK,$$
$$\dot{H} = s_H Y - dH$$

and:

$$A = A_0 e^{gt},$$
$$L = L_0 e^{nt}.$$

They analyze this model in exactly the same way. The regression equation that corresponds to the regression equation of the model without human capital looks like this:

$$\log\left(\frac{Y}{L}\right)_i = \bar{A} + \left(\frac{\alpha}{1-\alpha-\beta}\right)\log s_{K_i} - \left(\frac{\beta}{1-\alpha-\beta}\right)\left(\frac{\beta}{1-\alpha-\beta}\right)\log s_{H_i} - \left(\frac{\alpha+\beta}{1-\alpha-\beta}\right)\log(n_i+g+d) + \varepsilon_i,$$

where s_K is the fraction of output invested in physical capital, and s_H the fraction of output invested in human capital.

There are a number of problems. One problem is the following. We know how to get s_K out of the Heston-Summers data, because it is something that appears in the national accounts, but how do we get s_H? They use a variable that they call "school" which is the fraction of the labor force that is enrolled in secondary school.

This regression suggests a testable restriction: the sum of the second and third coefficient should just equal the negative of the third coefficient. When they estimate this regression the coefficients pass that test very well. The sum of the second and third regression coefficient minus the third regression coefficient is insignificantly different from zero and the significance level is 0.9. It does at least pass that test but the estimates that it gives for α and β are $\alpha = 0.4$ and $\beta = 0.3$ and therefore $(1 - \alpha - \beta) = 0.3$. This suggests the following production function:

$$K^{0.4} H^{0.3} (AL)^{0.3},$$

which fits the data quite well, and makes a lot of sense. It suggests that about half of the share of wages in total income is a return to "raw labor" and the other half is a return to human capital.

These authors find evidence of convergence of Y to predicted level after allowing for differences in s_H, s_K and initial Y. That is: use s_H, s_K to estimate $(Y/L)^*$. Then growth is faster for country i if:

$$\frac{Y_i}{L_i} < \frac{Y_i^*}{L_i^*}$$

in 1960.

Durlauf and Johnson (1992) find that the Mankiw, Romer and Weil data are better described if countries are divided into groups (roughly by per capita income): they find convergence within groups, but not between groups, and they find different production functions *between* groups.

One way to describe the whole situation is to say that these "tests" do not discriminate well between *endogenous-growth* models and standard exogenous growth models. The suggestion that about half of labor share is return to human capital sounds reasonable. Better measures of H would be useful.

REFERENCES

AGHION, P., HOWITT, P., 1992, "A Model of Growth Through Creative Destruction", in *Econometrica* 60, pp.323-351.

ARROW, K.J., 1962, "The Economic Implications of Learning by Doing", in *Review of Economic Studies* 29, pp.155-173.

BARRO, R.J., 1990, "Government Spending in a Simple Model of Endogenous Growth", in *Journal of Political Economy* 98, pp.S103-S125.

DE LONG, J.B., SUMMERS, L.H., 1991, "Equipment Investment and Economic Growth", in *Quarterly Journal of Economics* 106, pp.445-502.

DIXIT, A.K., STIGLITZ, J.E., 1977, "Monopolistic Competition and Optimum Product Diversity", in *American Economic Review* 67, pp.297-308.

DOMAR, E.V., 1946, "Capital Expansion, Rate of Growth and Employment", in *Econometrica* 14, pp.137-147.

DOMAR, E.V., 1957, *Essays in the Theory of Economic Growth*, Oxford University Press, Oxford.

DURLAUF, S.N., JOHNSON, P.A., 1992, "Local Versus Global Convergence Across National Economies", Mimeo, Stanford.

GROSSMAN, G., HELPMAN, E., 1991, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge, MA.

HESTON, A., SUMMERS, R., 1984, "Improved International Comparisons of Real Product and Its Composition: 1950-1980", in *Review of Income and Wealth* 30, pp.207-262.

HESTON, A., SUMMERS, R., 1988, "A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950-1985", in *Review of Income and Wealth* 34, pp-1-25.

HESTON, A., SUMMERS, R., 1991, "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988", in *Quarterly Journal of Economics* 106, pp.327-368.

JONES, C., 1992, "Time Series Tests of Endogenous Growth Models", Mimeo, MIT, March.

JONES, L., MANUELLI, R., 1990, "A Convex Model of Equilibrium Growth: Theory and Policy Implications", in *Journal of Political Economy* 98, pp.1008-1038.

KALDOR, N., 1957, "A Model of Economic Growth", in *Economic Journal* 67, pp.591-624.

KING, R.G., REBELO, S., 1990, "Public Policy and Economic Growth: Developing Neoclassical Implications", in *Journal of Political Economy* 98, pp.S126-S150.

LUCAS, R.E., 1988, "On the Mechanisms of Economic Development", in *Journal of Monetary Economics* 22, pp.3-42.

MANKIW, N.G., ROMER, D., WEIL, D.N., 1992, "A Contribution to the Empirics of Economic Growth", in *Quarterly Journal of Economics* 107, pp.407-437.

ROMER, P.M., 1986, "Increasing Returns and Long-Run Growth", in *Journal of Political Economy* 94, pp-1002-1037.

ROMER, P., M., 1990, "Endogenous Technological Change", in *Journal of Political Economy* 98, pp.71-102.

Nella stessa collana:

1) **GIUSEPPE DELLA TORRE**, Istituzioni creditizie ed Accumulazione del Capitale in Italia (1948-81), 1984.

2) **SANDRO GRONCHI**, Tasso Interno di Rendimento e Valutazione dei Progetti: una Analisi Teorica, 1984.

3) **ALDINO MONTI**, La Proprietà Immobiliare a Bologna in Età Napoleonica (1797-1810), 1984.

4) **GIULIO CIFARELLI**, Equilibrium and Disequilibrium Interpretations of Inflation, 1985.

5) **STEPHAN E. EPSTEIN**, Sicily and its markets: 1300-1500 economic growth and social transformation

Stampato nel 1992 dalla
COPISTERIA EMMECI, Via Pantaneto 132 - Siena -