NBER WORKING PAPER SERIES

HUMAN CAPITAL AND GROWTH: THEORY AND EVIDENCE

Paul M. Romer

Working Paper No. 3173

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1989

Prepared for the April 1989 Carnegie-Rochester Conference. The comments of the participants at the conference, the NBER conference on Growth, and the Applied Workshop at the University of Chicago, including comments made by Robert Barro, Gary Becker, Robert King, Paul Krugman, Alan Meltzer, Kevin Murphy, and Jim Poterba, are gratefully acknowledged. This work was supported by NSF Grant SES-8821943, a Sloan Foundation Fellowship, and the Center for Advanced Study in the Behavioral Sciences. This paper is part of NBER's research program in Growth. Any opinions expressed are those of the author not those of the National Bureau of Economic Research.

NBER Working Paper #3173
November 1989

HUMAN CAPITAL AND GROWTH: THEORY AND EVIDENCE

ABSTRACT

This paper outlines a theoretical framework for thinking about the role of human capital in a model of endogenous growth. The framework pays particular attention to two questions: What are the theoretical differences between intangibles like education and experience on the one hand, and knowledge or science on the other? and How do knowledge and science actually affect production? One implication derived from this framework is that the initial level of a variable like literacy may be important for understanding subsequent growth. This emphasis on the level of an input contrasts with the usual emphasis from growth accounting on rates of change of inputs. The principal empirical finding is that literacy has no additional explanatory power in a cross-country regression of growth rates on investment and other variables, but consistent with the model, the initial level of literacy does help predict the subsequent rate of investment, and indirectly, the rate of growth.

Paul M. Romer
Center for Advanced Studies in
the Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, CA 94305

1. Introduction

As suggested by its title, this paper offers both theory and evidence on the connection between human capital and growth, but only if both "theory" and "evidence" are interpreted broadly. The theory is really no more than a conceptual framework for thinking about growth, one that is intended to be useful in the analysis of data. It generates neither a set of equations to be solved for an equilibrium nor sharp quantitative predictions. Correspondingly, the empirical analysis presented here is no more than a preliminary attempt at exploratory data analysis guided by the framework.

Section 2 below presents the outline of the theoretical framework. Its conclusions for data analysis can be simply stated. If one allows for an explicit research and development activity designed to foster the creation of new goods, simple growth accounting relationships do not hold. In addition to the usual relationships between the <u>rate of growth</u> of inputs and the rate of growth of outputs suggested by growth accounting, there may be a role for the <u>level</u> of human capital variables to explain the rate of growth of output. In a regression equation that tries to estimate separate roles for both physical investment and human capital variables in explaining the rate of output growth, collinearity may cause the human capital variables not to enter significantly. They should still have explanatory power for investment.

The empirical part of the paper illustrates how this observation can be exploited by focusing on literacy as a measure of human capital. The results are consistent with what one would expect from the model (and also, one might add, from common sense), but they go beyond what one would expect from a narrow growth accounting framework.

The empirical analysis is also drawn into a discussion of the effects of measurement error. Formal tests for the presence of measurement error are relegated to a separate paper (Romer, 1989b), but the results here show that it is possible to correct for measurement error, at least in some cases and that doing so can significantly affect the inferences that one draws from the data.

Section 2, the bulk of the paper, outlines the theoretical framework in some detail. Because it is important for matching theory to data, the framework is explicit about exactly what one means in practice, i.e., in the available data, by education, experience, knowledge, and technology. Section 3 reports the results of cross country regressions. Section 4 summarizes the empirical findings and the contribution of the model.

2. Theory

2.1 Motivation

The usual approach in the study of growth is to outline a very specific dynamic model that can be explicitly solved for an equilibrium. In developing our sense of what happens in a new setting, explicit solutions are extremely important, but they are achieved at a substantial cost. Analytical tractability is decisive in the construction of such models, and artificial assumptions are inevitably made for purely technical reasons. As a result, when it comes time to compare the model with actual data, there is at best a distant and elastic connection between the variables manipulated in the model and those that we can actually measure. For example, I used a mongrel notion of aggregate capital that combines elements of both knowledge and physical capital (Romer, 1986.) It offers no clear guidance about whether physical capital, or physical capital plus cumulative research and development expenditures, or these two variables combined with expenditures on education and on the job training should be used in an empirical application of the model. Similarly, Lucas (1988) uses on a notion of human capital that grows without bound that apparently is quite different from the human capital measures like years of schooling and on the job training used by labor economists.

This section outlines an attempt at a model that lends itself more readily to a discussion of data. It builds on the model I have previously outlined (Romer, 1988), and extends that model's applicability by giving up any hope of deriving an explicit analytic solution. Based on the results that can be derived from the simpler model and other special cases of the general model, one can make informed conjectures about how the extended model will behave, but none of these conjectures is verified rigorously here. What this loose kind of framework can do is detail a list of possible variables to consider and a set of possible interactions to look for in an exploratory analysis of data.

Since the focus of this paper is education in particular and human capital more generally, the extension will focus on these variables and will be guided by the available data that bear on them. The empirical exercise in Section 3 will consider only the cross country data on literacy, but it will be clear that the analysis could be repeated with measures of primary, secondary, or higher education, or with measures of scientists, engineers, and technicians. To keep the discussion of the theory manageable, the model neglects the very important interactions between measures of human capital per capita and demographic variables like birth and death rates. It will also offer only a very simple specification of how the government interacts with the rest of the economy. For theoretical elaborations and empirical evidence on both of these points, see Barro (1989a).

2.2 Production of Goods

Let M denote the number of agents in a closed economy, and let i denote a typical individual. Each has a fixed allotment of time that can be divided between two different kinds of educational activities and four different productive activities. Every agent has an endowment of three types of skills:

L_i, physical skills like eye–hand coordination and strength;
E_i, educational skills acquired in primary and secondary school; and
S_i, scientific talent acquired in post-secondary education.

L_i will be taken as given, but it could be more explicitly modeled as the outcome of investments in nutrition, health care, and so forth.

The schooling measure E_i for each agent could be measured as it is in the data, in total years of schooling. Thus, for agent i, E_i grows according to

$$\dot{E}_i = \begin{cases} u_i^E & \text{if } E_i \leq 12, \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

where $u^E \in [0,1]$ denotes the fraction of time that is spent in primary and secondary school. (All rates of change will be denoted with an overdot, but nothing in what follows depends on the use of continuous time.) The total number of years of education in the population is then

$$E = \sum_{i=1}^{M} E_i, \tag{2}$$

the rate of growth of E will be

$$\dot{E} = \sum_{i=1}^{M} \dot{E}_i - \delta E, \tag{3}$$

where δ is the constant probability of death in any period. To keep the demographics simple, assume that one new person is born each time someone dies. Like many of the simplifying assumptions made here, the demographic assumptions could easily be made more realistic.

By convention, scientific skills S_i could be distinguished from skills acquired from primary and secondary schooling, and measured in years of post-secondary schooling. In some applications, one might choose a finer means of discriminating educational outcomes, distinguishing perhaps between college graduates generally and scientists, engineers and technicians. What matters here is only to illustrate how more than one type of skill might enter the production technology, and how different empirical measures of the more advanced skills could be used.

Corresponding to equations 1, 2, and 3 are equations describing how scientific skills evolve:

$$\dot{S}_i = \begin{cases} u_i^S & \text{if } E_i = 12, \\ 0 & \text{otherwise,} \end{cases} \qquad (4)$$

$$S = \sum_{i=1}^{M} S_i, \qquad (5)$$

$$\dot{S} = \sum_{i=1}^{M} \dot{S}_i - \delta S. \qquad (6)$$

As always in what follows, the variable u denotes the fraction of time devoted to an activity, so u^S denotes the faction of time devoted to scientific training.

Analogously, one can define a measure of cumulative job experience. Thus, let Z denote total man-hours of time spent on the job and specify that Z grows either through time spent working in the sector that produces consumption goods u_i^C or through time spent working to produce any of the intermediate inputs in production $u_i^{X_j}$:

$$\dot{Z}_i = u_i^L + \Sigma_j u_i^{X_j}, \qquad (7)$$

$$\dot{Z} = \sum_{i=1}^{M} \dot{Z}_i - \delta Z. \qquad (8)$$

This simple formulation is based on the assumption that experience is not too job specific or that job mobility is not too important. It could easily be extended.

It should be clear that allowing L and E and Z to enter as separate inputs into production is less restrictive than conventional specifications that assume that all three can be measured in common units of efficiency-man-hours. The specification here permits this as a special case, but does not require it. The inputs are kept separate in an attempt to be as explicit as possible about the different kinds of intangible inputs that are relevant for production. As the arguments that follow will show, an intangible input like education has theoretical properties that are very different from those of an intangible like an invention.

The specification here follows the convention from growth theory and aggregate general equilibrium theory of using as inputs in a production function only goods that have quantity units. This means that it does not follow the convention from labor economics of allowing some inputs to be measured in units that are the ratio of quantity units, e.g., years of education per worker. At a formal level, this choice is arbitrary, but for expositional purposes, the specification used here has distinct advantages. For example, in stating the neoclassical model with a Cobb-Douglas technology, one could specify output either as $Y = F(K,L) = K^{\alpha}L^{1-\alpha}$ or instead define $k = K/L$ and write $Y = G(k,L) = k^{\alpha}L$. These formulations are mathematically equivalent, but the second does not lend itself readily to general equilibrium analysis. The homogeneity of degree one of the true production function and its crucial role in the theory of distribution is hidden by the second formulation. The partial derivative $\frac{\partial G}{\partial L}$ has units goods/worker, but it is not equal to the wage rate. The partial derivative $\frac{\partial G}{\partial k}$ has units of goods/(unit of capital/worker) and has no interpretation as a wage or rental rate. The second formulation obscures much of the structure of the neoclassical model. An attempt to write the aggregate production function used here in terms of the number of workers and the average level of education and experience per worker would have exactly the same effect.

This framework highlights the fact that the variables E, S, and Z are bounded on a per capita basis. They cannot exceed the average length of life of people in this economy. For unbounded per capita income growth to take place, some input will have to grow per capita without bound. Average years of education or experience are not candidates for this variable.

It is clear that something like knowledge, understanding, or science has grown per capita and shows every prospect of continuing to do so. The very fact that unbounded growth is possible is an indication that this input is a very different kind of intangible from cognitive skill or memory related to the performance of a task. The feature that makes cognitive skill and memory easy to include in economic models is the one that makes them bounded: They are inextricably tied to a particular individual. An intangible like scientific or engineering knowledge is not tied to an individual. This means that they can grow without bound, but it also raises conceptual difficulties about the excludability and rivalry of knowledge as an economic good. These difficulties are discussed in more detail below.

Output of consumption goods in this economy will be denoted as C and expressed as a function of labor inputs $L^C = \Sigma_i u_i^C L_i$, educational inputs $E^C = \Sigma_i u_i^C E_i$, experience, $Z^C = \Sigma_i u_i^C Z_i$, and a list of intermediate producer durables $X^C = (X_1^C, X_2^C, \ldots)$. (Superscripts are used here to denote the productive activity that a particular input is used in.) Since u_i^C denotes the fraction of time agent i devotes to production of C, this person must supply all four of $u_i^C L_i$, $u_i^C E_i$, $u_i^C S_i$, and $u_i^C Z_i$ to this sector. By assumption, scientific skills make no contribution to increased output of C, so they are not reflected in the notation. The joint supply attributes of an individual's time, together with fixed time costs for acquiring educational and scientific skills and different relative productivities for the three factors in different sectors of this economy will lead to specialization in the acquisition of scientific skills. This issue is discussed by Becker and Murphy (1988), and is not pursued here.

In an abuse of notation, let $C(\cdot)$ denote the function that describes output of consumption goods as a function of its inputs: $C = C(L^C, E^C, Z^C, X^C)$. This formulation seems to leave out both knowledge or technology and capital. The omission of capital is only apparent. In the specification used here, different types of capital enter through the list of intermediate inputs X^C. A typical component of this list X_j could refer to the number of lathes, computers, or trucks.

The omission of knowledge or the technology is real. It reflects the belief that once the machinery X and the attributes of labor L, E, and Z have been specified, output is determined. Technology and knowledge enter production of consumption goods indirectly through their effects on the list of inputs X that is used at any point in time. In this sense, the model here is akin to models of technological change that is embodied in capital goods.

This formulation denies any role for science in producing E. A contrary claim that is sometimes made is that E years in school now produces a worker of higher productivity than E years in school did 100 years ago because of the growth of knowledge and science. Given that the main role of primary and secondary school education is to produce basic cognitive skills like the ability to read or to solve an equation for an unknown, this seems unlikely. School instruction today actually bears a remarkable resemblance to instruction 100 years ago. If there is any positive effect of science and technology on our ability to teach basic cognitive skills, it is small enough to neglect here. It is not through the schoolhouse that science has its effect on output. Rather, it is through the introduction of new goods.

To capture the process of new good introduction, the list X of actual and potential inputs is assumed to be of infinite length. At any time, only a finite number of X_j's, the ones that have been invented so far, can be used. For example, if X_ℓ denotes a computer that operates with pulses of light rather than with pulses of electricity, X_ℓ is now equal to 0 because no such computer is available yet. One can nevertheless conjecture how its availability would affect output, for example in fiber-optic networks, if it were. The simplifying assumption that the function C and the complete infinite list of arguments X_j is known from the beginning is not to be taken literally. The basic points of the analysis that follows will carry over to a model that introduces uncertainty about the function $C(\cdot)$ and the list X.

For a particular intermediate input of type j for which a design is already available, the stock of X_j will increase when production is undertaken. (Where designs come from is discussed below.) The stock will decrease with depreciation at a rate ρ. As for C, output of units of X_j can be written as a function of the amount of physical labor $L^{X_j} = \Sigma_i u_i^{X_j} L_i$, education skills $E^{X_j} = \Sigma_i u_i^{X_j} E_i$, experience $Z^{X_j} = \Sigma_i u_i^{X_j} Z_i$, and other producer durables X^{X_j} that are employed:

$$\dot{X}_j = X_j(L^{X_j}, E^{X_j}, X^{X_j}) - \rho X_j. \qquad (9)$$

Neither the stock of scientific knowledge nor that of scientific skills is a direct input of the manufacturing process for C or for the X_j's.

The conventional assumption in growth theory is that production functions like $C(\cdot)$ and $X_j(\cdot)$ are homogeneous of degree one. This assumption seems entirely reasonable in the present context. Most of the alleged scale economies in plant size or manufacturing processes should be exhausted at scales of operation that are small compared to the size of a national economy. Where a truly fundamental departure from the usual assumptions about returns to scale does arise is in the creation of a good, not in its subsequent production.

The essential fact about new good creation is that it requires expenditures that are quasi-fixed. These expenditures must be incurred to produce any goods at all, but they do not vary with the level of production. Such costs, typically subsumed under the label of research and development, include invention, construction of a prototype, testing, and refinement. Generically, these costs will be referred to as the cost of producing a design. The manufacturing function $X_j(\cdot)$ describes what happens when the design is sent to the factory floor for production.

Degree of excludability and rivalry are fundamental attributes of any economic good used in production or consumption. The observation that an intangible good like a design knowledge is often only partially excludable is very familiar. The fact that it nonrival has received less attention, but it is equally important. A design is a nonrival input in production in the sense that the same design can be used simultaneously in as many production processes as desired.

The extent of rivalry is determined entirely by the technology. In contrast, the notion of excludability is determined by both the technology and the legal institutions in a particular economy. If a good is purely rival, using it yourself is equivalent to excluding others from using it. If it is nonrival, excludability requires either a technological means for preventing access to the good (e.g. encryption) or a legal system that effectively deters others from using the input even though it is technologically possible to do so (e.g patents).

There have been repeated acknowledgments that the production of knowledge or technological change seems to conflict with the assumption of price-taking. Schumpeter (1942) gives one of the classic statements of the conflict between innovation and perfect competition. Arrow (1962a), Shell (1967), and Wilson (1975) do so as well. Despite the central role played by knowledge or technology in models of growth, growth theory has tended not to dwell on this issue. The neoclassical model of exogenous technological change (as presented by Solow, 1956) implicitly acknowledges the nonrival aspects of knowledge—improvements in the technology can be exploited simultaneously by all firms—but it does not consider the possibility that knowledge is privately provided. Arrow (1962b) allows for nonrival knowledge that is privately produced, but only as an unintended side effect of other activities. The formulation in my earlier paper (Romer, 1986) and in Lucas's paper (1988) introduces forms of knowledge that are partly excludable and rival, and partly nonexcludable and nonrival. As in the Arrow model, the nonrival knowledge is produced only as a side effect of some other activity. All three of these models allow for spillovers, that is, problems relating to excludability of knowledge, but they do not address the conflict between nonrivalry and private creation of a good in competitive markets.

Attempts to avoid the issue of the intentional, private production of nonrival inputs like designs or inventions presumably arise from technical considerations relating to the construction of economic models rather than a belief that privately produced nonrival goods are of negligible importance. Direct estimates of the magnitudes involved are not easy to come by, but we know that something on the order of 2% to 3% of GNP in industrialized countries is spent on research and development. Almost all of the output from this activity has the nonrival character of blueprints, designs, or inventions, and much of this activity takes place in the private sector.

A casual examination of the business press suggests that the problems for firms created by the private provision of a nonrival input are very real. For example, recent stories have described thefts of secret process technologies used by Du Pont in the production of Lycra, and of thefts of documents from Intel concerning its 80386 microprocessors. The problems in the micro chip and chemical industries have high visibility and are easy to understand, but large resources are at stake in more mundane areas like the design of blades for steam and gas turbines that are used to generate electricity. General Electric mounted extensive criminal and civil proceedings to keep its $200 million investment in mechanical drawings and metallurgical formulas for turbine blades from being used by competitors who had received copies of internal documents (Wall Street Journal, August 16, 1988, p.1.)

The point here is not just that it is costly to make knowledge excludable. Even if patent and copyright protection were perfect and completely subsidized by the government so that excludability was perfect, it is still the case that these firms, and many like them, sell goods that depend in an important way on nonrival inputs. Neither Intel, nor Du Pont, nor General Electric can sensibly be modeled as a price taker in the market for its goods. Once designs are in place, chips, Lycra, and turbine blades can all be manufactured under conditions of constant returns to scale. If these firms sold their goods at marginal cost, which is equal to unit cost, none of them would ever be able to recoup the initial research expenditures.

The nonrivalrous aspect of new good design is captured here by assuming that there is an additional variable A (A for "applied science") that represents the outcome of applied research and development. A separate knowledge input B (B for "basic science") will be introduced below, that is both nonrival and nonexcludable. The distinction between A and B arises from the fact that A is assumed also to be at least partly excludable, at least insofar as it is used in the production of a good. Thus one unit of A confers the right to produce a good that is protected by the legal system or secrecy from copying for at least some period of time. This means that it is possible to have the private provision of A.

Loosely speaking, A measures the total number of designs. One unit of A must be produced before it is possible to start production of each new good X_j, so A is also a measure of the total number of types of goods that are available. Thus $X_j = 0$ for any $j>A$. For $j \leq A$, the level of production of X_j depends only on the inputs used in production, $X_j(L^{X_j}, E^{X_j}, Z^{X_j}, X^{X_j})$.

2.3 A Simple Special Case

In its general form, this model is like an infinite extension of a model with multiple capital goods. A special case that permits a significant simplification and still captures the essential issues about the role of science and knowledge is one where the production functions for manufacturing capital goods and consumption goods have the same functional form:

$$X_j(L,E,Z,X) = C(L,E,Z,X) = G(L,E,Z,X), \text{ for all } j = 1, 2, \ldots (10)$$

This assumption suppresses any factor intensity differences in the different production activities. The inputs L, E, Z, and X will be used in identical ratios in all of the productive activities, and a change in the output of one activity requires only a scale change in the other activities. Using the kind of functional form suggested by Dixit and Stiglitz (1977) and Ethier (1982) for modeling the dependence of utility or production respectively, on many goods, a simple example of a functional form for $G(\cdot)$ is

$$G(L,E,Z,X) = L^{\alpha}E^{\beta}Z^{\gamma}\sum_{j=1}^{\infty} X_j^{\mu}. \qquad (11)$$

Homogeneity of degree one is imposed by the assumption that $\alpha+\beta+\gamma+\mu=1$.

A further simplification illustrated by this kind of function is the assumption that all of the different types of producer durables X_j have symmetric effects in production. With both of these assumptions, it is possible to define aggregate capital as $K = \Sigma_j X_j$. Because of the symmetry assumption, the concavity of the function G implies that all of the inputs X_j that are available will be used at the same level, $X_j = X_k$ for all j and k less than or equal to A.

With these assumptions, it is possible to define aggregate output Y as a function of total labor used in production of goods, $L^Y=L^C+\Sigma_j L^{X_j}$; total educational inputs used in production of goods, $E^Y = E^C+\Sigma_j E^{X_j}$; total years of experience in the labor force, $Z^Y = Z^C+\Sigma_j Z^{X_j}$; total capital used in the production of goods, $K^Y = \Sigma_k \Sigma_j X_j^{X_k} + \Sigma_j X_j^Y$; and the number of designs or goods in existence, A. Even though A is not a true input in the production of goods and does not appear in any of the manufacturing functions $C(\cdot)$ or $X_j(\cdot)$, its presence is crucial to distinguish between a case in which an increase in K is caused by an increase in the quantities of existing intermediate producer durables and a case in which the increase in K is caused by an increase in the number of types of intermediate producer durables.

Let $X(K^Y,A)$ denote the list of producer durables with the property that $X_j = K^Y/A$ for $1 \leq j \leq A$ and $X_j = 0$ for $j > A$. Then define

$$F(L^Y,E^Y,Z^Y,K^Y,A) = G(L^Y,E^Y,Z^Y,X(K^Y,A)). \qquad (12)$$

For the functional form for $G(\cdot)$ given in equation (11), $F(\cdot)$ takes the form

$$F(L,E,K,A) = L^{\alpha}E^{\beta}Z^{\gamma}K^{\mu}A^{1-\mu}. \qquad (13)$$

At any time, $Y = F(L^Y,E^Y,Z^{\gamma},K^Y,A)$ represents the total feasible output that can be split between consumption and accumulation of additional inputs X, or equivalently of additional K:

$$Y = F(L^Y,E^Y,Z^Y,K^Y,A) = C+\dot{K}. \qquad (14)$$

The simplifying assumptions and functional form described here are not essential for the arguments that follow, but the resulting expressions (12), (13), and (14) do help make concrete some of the abstract arguments and show the close parallel with earlier models. For fixed A, the model here reduces to a description of the technology to one that is very close to the one assumed in the standard neoclassical model. Physical labor, education, and experience are entered separately to avoid some of the ambiguity associated with the use of the term human capital, but otherwise the specification is entirely conventional. In particular, it follows that F is homogeneous of degree one in its first four arguments, holding A constant, just as it is in the neoclassical model.

The theoretical tasks that remain are to specify the technology for producing new designs, and more important, to specify an equilibrium that supports production of A by private firms. For these tasks, the entire specification leading up to equation (14) is needed, not just the reduced form equation itself.

2.4 Production of Designs

The production technology for creating new designs measured by A is assumed to depend on the amount of scientific and educated labor S^A and E^A used in this process. It depends on the amounts of intermediate inputs X^A, for example computers, used for this purpose. It is also assumed to depend on the stock of basic science B that is known. Finally, the production of new designs is assumed to depend directly on the stock of existing designs because existing designs offer hints about how to undertake future designs. Thus, for example, the productivity of an engineer with 8 years of post-graduate training who is engaged in the design of a new good will depend on both of the measures A and B of cumulative knowledge available for use.

$$\dot{A} = \dot{A}(E^A,S^A,A^A,B^A,X^A), \qquad (15)$$

It is tempting to assume that A suffers depreciation as well. After all, things like designs or blueprints do sometimes get lost. However, this is not what one typically has in mind here. Instead, depreciation is used as a synonym for obsolescence, but the distinction here is important. A particular piece of engineering may lose its economic value without being truly lost. The canonical example is the design for a buggy whip. This kind of obsolescence is not explicitly captured in the model as it stands. To do so, one would have to modify the simply functional form used in equation (11), in which all intermediate inputs enter production in an additively separable way, and allow instead for complicated patterns of complementarity (say between buggies and whips) and substitution (between horse power and internal combustion.) This extension is of real interest, but adding an exponential depreciation term to equation (15) seems like a poor, and possibly misleading, substitute for it. In any case, the basic lesson here is likely to survive any extension: growth accounting based on perfect competition leads to too narrow a focus on rates of change of inputs.

The production of basic science depends on the amount of scientific talent S^B devoted to this activity, its own level B, and any of the intermediate inputs X that are available for use:

$$\dot{B} = \dot{B}(S^B, B^B, X^B). \qquad (9)$$

Once one constructs this kind of explicit framework, it is clear how it could be extended. For example, to model learning by doing, arguments of the production function for C or for the X_j's could also appear as arguments in the production of A. For example, if people on the job in the production of Y have insights about new products or processes purely by virtue of doing their jobs, time spent on the job L^Y (or L^Y and the educational level E^Y) would appear as arguments of $\dot{A}$.

The constraints on the rival inputs in this model are entirely conventional. The constraint on the allocation of time is

$$u^C + \Sigma_j u^{X_j} + u^E + u^S + u^A + u^B \leq 1. \tag{17}$$

This constraint determines the allocation of the inputs L, E, and S to the various production activities. Similarly, the stock of X must be allocated among production of Y, A, and B. The constraints on the nonrivalrous goods are different precisely because they are nonrival:

$$\begin{aligned} A^A \leq A,\ A^B \leq A, \\ B^A \leq B,\ B^B \leq B. \end{aligned} \tag{18}$$

It is possible that these last constraints are not met with equality. If part of A or B developed by one organization is kept secret, it may not be used in subsequent production of A or B by other organizations.

There are important questions about aggregation that are not being addressed here, but it should also be clear how they could be addressed. Output of both A and B could be indexed by the producing organization, with individually indexed levels of inputs. Total output would be the sum of across firms corrected for double counting (i.e. for the production of the same piece of A or B by different firms or labs).

2.5 Prices and Marginal Products

At the level of generality used here, there is not much that one can prove rigorously about this system of equations. One immediate implication of the presence of nonrival inputs in production, however, is that the competitive assumptions needed for a complete accounting for growth do not hold. At the firm level, the failure of the usual assumptions for competition follows from the decreasing average total cost of producing X_j implied by the initial investment in design costs. If the firm priced output at marginal cost (equal to constant unit cost) as competition would force it to do, it would never recoup this initial investment.

At the aggregate level, this departure from the usual assumptions shows up in the form of aggregate increasing returns to scale. Consider an economy that starts from initial stocks L_0, E_0, S_0, Z_0, K_0, A_0, B_0 and evolves through time. If the economy were instead to start with twice as much of the initial tangible stocks L_0, E_0, Z_0, S_0, K_0, it would be possible to produce more than twice as much consumption good output at every point in time. It could produce exactly twice as much by building a second economy that replicates the production of the rivalrous goods C and all of the X_j's and that replicates the accumulation of E, Z, and S that takes place in the first economy. Since the underlying production functions for C and X_j are homogeneous of degree one, as are the schooling technologies, this is feasible. At every point in time, this replica economy could make use of the stock of the nonrivalrous goods A and B that was already available in the original economy. Even if the portion of the talent E and S and of the inputs X that are used to increase A and B in the first economy is left idle in the replica economy, it can replicate all of the output of the first economy. If the idle E, Z, S, and X were instead used to produce additional units of A or B or merely used in production of C or of the X_j's, output would more than double. Thus aggregate output increases more than proportionally with increases in the rivalrous inputs L, E, Z, S, and K alone. Once one recognizes that A and B are inputs too, measured, say, in units of production cost, it is clear that a proportional increase in L, E, S, K, A, and B would increase output by even more.

The fact that it is not possible to replicate any number of existing inputs is not relevant here. All that matters in this thought experiment is what it can reveal about the underlying mathematical properties of production. Because of the departure from homogeneity of degree one that it reveals, it follows that market prices cannot reflect marginal values.

In a simple static model, a production function that increases more than proportionally with increases in all of the inputs has the property that the marginal product of each input times the quantity of that input, summed over all inputs, yields a quantity that is greater than output. A marginal productivity theory of distribution fails because paying each input its marginal product would more than exhaust total output. In a dynamic model, this result is repeated at every point in time.

2.6 Equilibrium

The discussion so far has established the basic elements of the technology and indicated why a classical competitive equilibrium will not be feasible. This section proposes an alternative equilibrium concept.

The simplest alternative would be to assume that no compensation is paid to any resources that are devoted to the production of A or B. In this case, all output $Y = F(L^Y,E^Y,Z^Y,K^Y,A)$ at given time can be distributed as payments to L, E, and K according to marginal productivity theory because F is homogeneous of degree one in its first four arguments. This is the alternative followed by the neoclassical model with exogenous technological change. The problem of course is that since the output sector pays nothing for the use of the designs A, there is no way to compensate the inputs used to produce more designs A and more basic science B, so these inputs are ignored or suppressed. The spillover models follow this same route if one makes the qualification that the relevant marginal products for distribution theory are private ones, not social ones. Inputs used intentionally to produce A or B are still suppressed.

Models that rely on government funding to pay for increases in A and B reintroduce these inputs. In effect, the power of the government to tax is used to break the budget constraint on overall resources. All of GNP can be paid to the factors other than A in competitive markets according to marginal productivity theory. Then some of this income can be taken away by means of lump sum taxation and used to compensate inputs used to produce increments to A and B.

The essential characteristic of basic research is that it is very close to a purely nonexcludable, purely nonrival good; that is, it is a classical public good. It is very hard to establish any kind of property right over basic scientific discoveries. Consequently, private firms engage in very little true basic research unless they are paid to do so by the government. In developed countries, government funding probably gives a reasonable description of the process that leads to growth in basic science. In less developed countries, which can take advantage of the results produced in the developed countries, growth in basic science is essentially an exogenously given feature of the world that they inhabit.

If basic research were the only kind of nonrival good, this is as far as the theory would need to go. The neoclassical model of exogenous technological change, supplemented in developed countries by a model of government subsidies for basic research, would be sufficient to understand growth. Empirically, growth accounting would need only to be supplemented by an analysis of government support for basic science, possibly worldwide.

Appealing as this description is at a theoretical level, it is not very useful empirically. Part of the point of the detailed outline of production give above is that basic research is a not a direct input in the production goods. Before it can be exploited, someone has to design and develop a good that takes advantage of it. A very large part of this kind of activity takes place in the private sector, and most of it is undertaken as an intentional investment activity. The spillover models are surely correct in the sense that firms rarely capture all of the benefits created when they undertake research, but these models fail to explain the underlying motivation for doing the research in the first place. They do not explain why the knowledge produced from intentional private–sector research and development generates a return.

The most natural model explanation of how this can happen, and the one emphasized by Schumpeter, is through the presence of market power. Research and development leads to a new good that is not a perfect substitute for any existing good in the market. The producer of the good can exploit the unique qualities of the good and sell it at a price that is greater than its unit cost of production. This observation must be combined with the observation that introducing new goods is subject to free entry. In equilibrium, all firms must earn zero profits in the relevant sense: the initial cost of designing and developing a good must be just equal to the present discounted value of the difference between the unit cost of producing the good and the price that the firm charges for it. Each good X_j can be thought of as being introduced by a different firm. The resulting equilibrium is one with monopolistic competition between a large number of firms engaged in the introduction and production of new goods X_j.

It would clearly be possible to introduce fixed costs and departures from price-taking behavior in sectors other than the intermediate good sector. As it stands, the model assumes that there is a single final good that is produced according to a production function that is homogeneous of degree one, which leads to price taking in the final goods sector. One could easily allow for the possibility that there are also many distinct consumption goods supplied under conditions of monopolistic competition. For the purposes of this paper, this would add little. Thus, for thinking about equilibrium, all goods other than the intermediate inputs are assumed to be sold in competitive markets.

2.7 Empirical Implications

For a representative country, the implications of this framework can be elicited from Figure 1. As noted above, the rate of growth of basic science B for most purposes can be taken to be exogenously given by developments in other countries and by government policy decisions. This does not reduce the model to the neoclassical model, because B has no direct effect on measured output Y. What matters for growth in output is growth in A. If A is constant because no resources are devoted to converting the stock of B into usable products, the fact that education and experience are bounded per capita implies that the economy will reach a conventional stationary state in per capita income because of diminishing returns to the accumulation of K; that is, because of the diminishing marginal productivity of more units of the same set of types of producer durables.

Figure 1 illustrates what happens to conventional growth accounting relationships in an equilibrium in which A grows, i.e., one in which new goods are constantly being introduced. It plots an illustrative graph of total output Y as a function of the amount of a specific intermediate input X_J when other inputs are held constant. Let P_J denote the rental rate charged for this input by the single firm that produces it. For simplicity, assume that once the good is introduced, the rental rate is constant for all time. Since firms that buy this input are price takers, the durable good will be used at a level $\tilde{X}_J$ such that its marginal productivity is equal to P_J. Under the conditions of zero profit that must obtain under monopolistic competition, revenue to the seller $P_J\tilde{X}_J$ is equal to the total cost to society of supplying good J at the level $\tilde{X}_J$. For the present value of entry to be zero, the revenues at any given time $P_J\tilde{X}_J$ must be just equal to the interest cost on the initial investment to design the good, plus the sum of the interest rate and the depreciation rate times the cost of producing the $\tilde{X}_J$.

Because of the concavity of the production function, the introduction of a new intermediate input generates an additional flow of consumption goods at every point in time that is greater than the cost, measured in consumption goods, of the resources that are used to produce the intermediate input. In the figure, this follows from the fact that the intercept of the tangent to the curve is greater than zero. This intercept represents the increment in the value of output that cannot be accounted for by an increase in the value of the raw inputs L, E, Z, and K used in production.

(A similar point could be made about the introduction of new consumer goods instead of new intermediate inputs. Increases in GNP will understate increases in welfare when new goods are introduced because expenditure on new goods does not take account of the additional consumer surplus added by the good. Since welfare is not measured, however, this effect has no obvious implications for the analysis of cross country data on growth.)

The most direct implication of this kind of model is that data on something like patents or new good introductions, or even private applied research and development spending, would have important explanatory power for growth. As a practical matter, internationally comparable data of this kind are not available for many countries. Even among the developed countries that try to maintain aggregate statistics on research and development, there are likely to be important differences in the conventions for defining an activity like new good development and measuring expenditure on it.

An indirect strategy for explaining cross-country variation in the growth of A is to focus on the inputs that determine its rate of growth. One obvious input is the rate of growth of basic science B, but it is unlikely that this can be exploited in a cross-section regression. By assumption, basic science is a nonexcludable good and can be exploited anywhere in the world once it is produced.

Another variable that influences the rate of growth of A should be the level of A in a country relative to that in the rest of the world. A country that exploits a very small range of inputs relative to the range of inputs available in the rest of the world might be expected to be capable of rapid growth in A, either through the initiation of international trade or through copying and reverse engineering. As noted above, direct measures of A or its rate of change are not available, but it is likely that a low level of A is associated with low income per capita. Thus, if everything else is held constant, one might expect that poor countries would tend to have faster growth in A, and therefore faster growth in output. This is a version of the catching-up hypothesis also generated by the neoclassical model, but here the process in not automatic. The inputs that one would expect to affect the rate of growth of A are openness to trade and measures of total educational achievement and of scientific talent in a country. It is not logically necessary that a country with more educational talent actually devotes it to new good production, but the presumption is that it would, and in at least one worked-out model (Romer 1988) this is the case. This paper also contains a theoretical demonstration that free trade should increase growth. Empirical consideration of trade variables is put off for future work. Here, only measures related to the educational variables are considered in the illustrative empirical analysis in the next section.

A special case of the model outlined here that has been solved (Romer, 1988) combines the variables E, Z, and S into a single human capital variable H and combines basic research B and applied product development A into a single variable A. The symmetry assumptions from Section 2.2 are used, together with a functional like that given in equation (11) for the functions that depend on the infinite list X. With these specifications, the model permits an explicit solution for a balanced growth equilibrium. Increases in the total stock of education and scientific talent lead to increases in the amount that is allocated to the production of A. Generalizing to the model here, one should expect that the rate of growth of

A is an increasing function of the level of E and S in the economy. The rate of growth of A should in turn help explain the rate of growth of K and the rate of growth of income.

In the balanced growth solution calculated for the special case, the rate of growth of A is identical to the rate of growth of K. New investment takes place one for one with growth in the new opportunities represented by growth in A. Thus, in a regression of the rate of growth of output on investment and the level of education and scientific talent, collinearity between investment and $\dot{A}$ would mean that there is nothing left for the level of education and scientific measures to explain. They are proxies for growth of A, but investment is an even better proxy. Investment will have a significant partial correlation with the rate of growth, in contrast to the neoclassical prediction that in the steady state, the rate of investment should not be associated with with the rate of growth. This association arises because investment picks up the indirect effects of increases in A. In this case, one would expect that a regression of investment on the educational variables should reveal an important partial correlation.

In summary, the empirical implications of this analysis are that the *level* of a human capital variable like education or scientific talent will be correlated with both the *rate* of growth of income per capita and the share of total output devoted to investment in physical capital. It is possible that the educational variables will not be significant in a regression for output growth that also includes the rate of investment. If so, the rate of investment should be significantly related to growth, even in the long run, and the rate of investment should be related to the level of education. Finally, a tendency for less developed countries to catch up because of more rapid growth of A should lead to a negative partial correlation between the initial level of income and the subsequent rate of growth.

A prediction that this model shares with more conventional models is that the rate of growth of human capital variables should also be related to the rate of growth of output. This is the conventional presumption from growth accounting. Like the rate of growth of A, variation across different countries in the steady state the rate of growth of education may tend to be matched by variation in the rate of growth of investment. Thus in a regression context, growth in human capital may not be significant in a regression of growth rates on a list of variables that includes investment. Once again, investment should explain the rate of growth of output, and the change in human capital should explain investment.

This model does not offer any direct test of the proposition that there are increasing returns at the aggregate level. As the logic of the model shows, presence follows from the assertion that nonrival goods like inventions, designs, or science are important for explaining long run growth. That this is true is suggested by industry studies of the productivity of research and development, but comparable evidence at the aggregate level will have to await the construction of measures of research and development that are comparable across countries.

Section 3. Empirical Results

3.1 Description of the Data and Related Work

The basic source of national income accounts data used here is the World

Data table compiled by Robert Summers and Alan Heston (1988). The measures of human capital collected come from the United Nations, primarily from the annual statistical yearbooks published by UNESCO. These include obvious measures like literacy and less obvious measures like consumption of newsprint per capita and the number of radios per capita. To keep the project manageable and because of data limitations, consideration of measures of higher level human capital like the number of college graduates of the number of scientists and engineers is put off for subsequent work, and the results presented here merely illustrate the kind of data analysis that could be done. The results reported here are concerned only with the connection between basic literacy and the rate of growth of income per capita and the rate of investment. Literacy was chosen partly because it is a variable that is available for a broad sample of countries, and partly because cross-country measures of literacy should be more comparable than cross-country measures of educational attainment. Finally, literacy has the advantage that its level at any time is easily measured. This is useful for the analysis undertaken below that differentiates between the level of a human capital variable and its rate of change. Most of the data on education are flow measures like enrollment rates rather than stock measures of average educational attainment.

I used data from an earlier version of the world data table constructed by Summers and Heston were used in a preliminary investigation of cross country variation in per capita growth rates and investment (Romer, 1987, 1989a.) These data have also been used in conjunction with detailed data on government expenditure and demographic variables by Barro (1989a) in an analysis that focuses on fertility choice and on a possible productive role for government investment expenditure, and with emphasis on human capital (Barro, 1989b). His results are not strictly comparable with results here.

Barro's estimates make use of variables that are not used here. In addition, because of the limited availability of data for some variables, the sample of countries considered differs. This problem recurs throughout all of the subsequent analysis. Anytime a variable other than one from the Summers and Heston data set is added to a regression, the number of countries with complete data gets smaller. Other than Barro's, the work most closely related to the results reported here are the regressions reported by Azariadis and Drazen (1988), who also regress cross-country measures of growth rates on literacy. The results reported here are qualitatively similar to theirs, but the interpretation here raises the possibility that measurement error plays a role in their findings.

There is of course a very large literature on human capital generally, and human capital as it relates to growth accounting (so large in fact that it is a challenge for a nonspecialist to read even the surveys in the area.) Without making any attempt to give a balanced overview of this literature, I can offer some subjective impressions. There is lots of evidence that across individuals, the level of education is correlated with all kinds of indicators of ability and achievement. Because economics is not an experimental science, it is not easy to draw firm conclusions about the causal role of increases in education on earnings at the individual level or on output at the aggregate level. Probably the strongest evidence is the general finding that agricultural productivity is positively correlated with the level of education of the farmer (see for example Jamison and Lau, 1982.) This evidence has the advantage that farmers are generally self-employed, so that signaling is not an important issue, and inputs and outputs can be measured relatively directly.

This evidence leaves open the possibility that unmeasured attributes cause both the variation in educational achievement across individuals and the variation in productivity, but there is separate evidence such as that presented by Chamberlain and Griliches (1974, 1979) using sibling data on education, labor market outcomes, and test scores to suggest that unobserved attributes are not so large as to overturn the basic finding that improvements in education cause improvement in economic outcomes.

Taken together, the accumulated evidence suggests that education almost surely has a causal role that is positive, but beyond that our knowledge is still uncomfortably imprecise. Moreover, there seems to be a general sense that the "human capital revolution" in development has been a disappointment, and that growth accounting measures of the effects of education do not help us understand much of the variation in growth rates observed in the world. In this context, one of the questions that this particular exercise faces is whether different theory and the use of different ways of looking at the evidence will increase our estimate of the empirical relevance of education for understanding growth. From this point of view, the results here are mildly encouraging. Previous analyses may have missed an important channel whereby education has fostered growth.

3.2 Regression Results

The list of variables considered is presented in Table 1, together with basic descriptive statistics. The sample of countries used in initial investigations included all of the market economies from the Summers and Heston data set for which data are available for the entire period from 1960 to 1985. The initial plan was to retain all of the high-income oil-exporting countries (as defined by the World Bank), but to allow a dummy variable for countries in this class. Much of the subsequent analysis turns on the properties of the initial level of real income per capita in 1960, however, and at roughly $50,000 (in 1980 dollars) Kuwait is an outlier by an order of magnitude. The next highest value is for the U.S. at around $7000. Moreover, of the high-income exporters, only Kuwait and Saudi Arabia had enough data to be included in the sample. So that Kuwait would dominate all of the regressions, it was excluded. Since Saudi Arabia was the single remaining high-income oil-exporter, it too was dropped. The remaining sample consists of 112 countries.

The starting point for the analysis is the regression described in Table 2. It gives least squares estimates of the effects that the initial level of income, the average share of total investment (including government investment) in GDP over the sample period, and the level of literacy in 1960 have on growth of income per capita from 1960 to 1985. For comparability with other other results, the regression also includes the level of government non-investment spending and dummy variables for the continents of Africa and Latin America (including Central America and Mexico).

Because of the relatively long time period (25 years) over which rates of growth and the average share of investment in GDP are measured, the fixed weight price series from Heston and Summers are avoided. The real income series used to calculate the rate of growth and the level is a chain index. The share variables are measured in current prices. The qualitative results are not too sensitive to this choice. Changes in parameter estimates due to the use of different measures of income and share variables are not large compared to the estimated standard errors.

Table 2 reports the results of a least squares regression of the rate of growth on the level of income, Y60, current price measures of the share of investment, INV, the share of government consumption spending, GOV, dummy variables for the continents of Africa and Latin America, AFDUM and LADUM, and literacy in 1960, LT60. Attempts to uncover heteroscedasticity, some of which were reported in an earlier version of this paper, found some evidence that the poorest countries tended to have somewhat larger residuals; even so Generalized Least Squares estimates did not lead to important changes in any of the coefficient estimates or inferences, and they are not reported here.

The basic results in Table 2 seem to be consistent with the hypotheses suggested above. In particular, the initial level of income has a significant negative partial correlation with the rate of growth and the initial level of literacy has a positive partial correlation. Table 3 reports a test of the robustness of these findings that was motivated by a concern that there could be important measurement error in the initial level of income and in the reported literacy rate. The possibility of measurement error in the level of income is particularly important because it is used to calculate the growth rate on the left-hand side. Any measurement error in this variable will induce a spurious negative correlation between growth and the initial level of income. Since the level of income is also closely correlated with the level of literacy, bias in the coefficient for the level of income is likely to cause a bias of the opposite sign in the estimate of literacy. (For a proof that this is the case and further discussion of the problem of measurement error, see Romer 1989b.)

Table 3 suggests that this is exactly what happens. Once instrumental variables are used to correct for measurement error in both the initial level of income and the literacy rate, neither is significantly related to the rate of output growth. In addition to the other variables used in the equation, the instruments used here include the log of newsprint consumption per capita in 1960, NP60, and the number of radios per 1000 inhabitants in 1960, RD60. Newsprint consumption was expected to be a useful indicator of true initial literacy. Because the distribution of values for consumption of newsprint per capita turns out to be very significantly skewed, the logarithm of the level per capita was used as the instrument in the equations. Compared to the raw level, it improves the fit in the first stage regression of literacy on the instruments. The number of radios per capita was expected to act primarily as an indicator of income per capita.

Recall that for a variable to be a valid instrument, it need not itself be free of error or be perfectly correlated with the variables of interest. All that is required for a variable to be a valid instrument is that it be correlated with the variables used in the equation and that it be uncorrelated with the error term in the equation. Since these variables are both predetermined, they are good candidates for satisfying this last requirement.

To verify that the instruments are indeed correlated with literacy and the level of income, one need only consider the first stage regressions. The regression of literacy on the other variables has an R^2 of .78, with virtually all of the explanatory power coming from the newsprint consumption variable (t statistic of 6). The regression of the initial level of income has an R^2 of .84, with most of the explanatory power coming from the radios variable (t statistic of 8), and some from the newsprint variable (t statistic of 4).

From these results, one can conclude that the negative results reported in Table 3 are not simply the result of a bad choice of instruments. Rather, the results suggest that the apparently positive findings in Table 2 are attributable to measurement error. Because literacy and the level of income are correlated, and become more so once one corrects for measurement error, the problem here is partly one of multicollinearity. If one is willing to assume that one variable has no effect, it is possible to test for an effect for the other. Excluding the level of income, there is no evidence that the level of literacy has a separate effect in a regression that includes investment, whether or not one corrects for measurement error; but excluding literacy, there is some evidence (e.g., a t statistic of 1.8) that the level of income is negatively related to the rate of growth, even after correcting for possible measurement error.

Elsewhere, (Romer, 1998b) I report results showing that it is possible to reject the hypothesis that there is no measurement error in the level of income at conventional significance levels. Overall, the evidence for measurement error in the literacy rate is mixed. One cannot reject the hypothesis of no measurement error at conventional significance levels, but the estimate of the coefficient tends to change by a nontrivial amount in the direction that one would predict if measurement error were present. The use of the instrumental variables estimator entails at most a loss in efficiency, a loss that in this case appears to be acceptable. Hence, for the rest of the regressions, instrumental variables are used not only for the level of income but also for the literacy rate in 1960.

Contrary to what one would expect from the basic prediction of steady state dynamics in the neoclassical model with the same rate of technological change in all countries, the rate of investment is significantly correlated with the rate of growth in both Tables 2 and 3. Across papers, data sets, and specifications, this finding is the most robust partial correlation in cross-country data. This could either be an indication of a direct causal role for investment in causing growth in A (as was argued in Romer 1986, 1987), or of causality that runs from more rapid growth in A to higher rates of investment (as suggested here), or of more complicated forms of joint causality.

Like the coefficient on investment, the coefficients on the other variables in Tables 2 and 3 are not particularly sensitive to the use of instrumental variables. Since they are not considered by the theory, results concerning them will only be noted briefly. Consistent with the findings of others (e.g. Barro 1989a), the dummy variables for Africa and Latin America are significant and negative. This suggests that there are important omitted variables here. The government spending variable is negative, a result that is suggestive of inefficient intervention or tax effects that act through some mechanism other than through reductions in the rate of investment. For further evidence on the effects of the government, see Barro (1989a).

Both standard growth accounting and the theory suggested here predict that the change in the level of literacy between 1960 and 1980 should influence the rate of growth. This variable does not have significant effects if it is added to either of the specifications in Tables 2 or 3. Barro (1989b) reports similar results using primary and secondary school enrollment rates in 1960 as a proxy for growth in human capital. In both cases, this finding may reflect a problem of collinearity with the rate of investment like that as described for growth in A. In a model of steady state or balanced growth, both growth in the technology A and growth in the quality of the labor force would be expected to cause growth in K.

Table 4 presents evidence bearing on the conjecture that that literacy does not have a significant coefficient in Table 3 because investment is included in the regression. Consistent with this view, the table shows that <u>level</u> of literacy has a significantly positive partial correlation with the rate of output growth in an equation that excludes the rate of investment.

The <u>change</u> in literacy does not, however, have any independent effect on the rate of growth. In contrast, using enrollment data in 1960 as a proxy for the rate of growth of human capital, Barro (1989b) finds that they do have a significant positive partial correlation with the investment share.

Table 5 presents a more direct test of the assertion that the initial level of literacy and its rate of change are positively correlated with the rate of investment. It reports instrumental variables estimates of the coefficients in a regression of the share of investment on the initial level of income and literacy, on the change in literacy (LT_DIFF) and on the government share variable and the continent dummies. The literacy variable has effects that are both statistically and economically significant. An increase in the literacy rate from the mean of 50% to 60% is associated with an increase of the share of investment in GDP from the mean of 14% to 16%.

In this equation, it is reassuring for both this theory and conventional growth accounting that the <u>change</u> in literacy also exhibits a significant partial correlation with the rate of investment. This finding is consistent with the view that exogenous increases in literacy, hence in human capital, could cause increases in output that are collinear with increases in investment. Nevertheless, one must be especially careful in interpreting this partial correlation. It is quite possible that investment, growth in income, and growth in literacy are simultaneously determined. It is also possible that there is measurement error in the change in the literacy rate, but if the standard used to measure literacy is constant within a country over time, measurement error in the change may be less severe that measurement error in the level that is caused by cross-country variation in the standard. In contrast to the variables that are predetermined, for which there are good instruments, there are no good candidate instruments for the change in literacy, and the question of bias in its estimated coefficient cannot be resolved.

4. Conclusion

As suggested in the beginning, the empirical work undertaken here is more in the spirit of exploratory data analysis than of hypothesis testing. The role of the theory is not so much to generate sharp testable predictions, but rather to guide the process of data analysis. At best, the results demonstrate a weak consistency between the data and the theoretical framework outlined in Section 2. They should not be taken as a strong confirmation of the model considered here because there is little reason to believe that these data discriminate strongly between different models. The initial level of literacy helps explain investment, but it is correlated with other variables like life expectancy at birth, and an equation like that reported in Table 5 cannot discriminate between the hypothesis that it is actually life expectancy that causes higher investment rather than higher literacy, or that some other variable causes all three to move together.

Consistent with other similar empirical exercises, the regressions here report a strong and robust finding that the rate of investment helps explain the rate of growth in cross-country regressions. After one corrects for measurement error, the results here find only weak evidence that the initial level of income is negatively related to the rate of growth. The only measure of the growth rate of a human capital variable considered here, the change in the level of literacy, does not have an independent effect in the growth equation, but it does help explain the rate of investment, as one would expect from a steady-state version of the model here or of a conventional neoclassical model.

The interesting, and from the point of view of conventional theory unexpected finding is that the initial level of literacy also helps explain the subsequent rate of investment, and thereby the subsequent rate of income growth. It suggests a direction for further work that would not have been evident from rigid adherence to a conventional neoclassical model or its empirical counterpart, growth accounting.

The substantive contribution of this paper is not to offer a firm set of conclusions about the causal relations underlying the data, but rather to demonstrate how a more general model can be used to guide the analysis of data. The model suggests the inclusion of a variable that might otherwise have been neglected, and it turns out to be important.

Other models might have suggested this variable, and as suggested in the introduction, even common sense might have done so. Nonetheless, one should not underestimate the practical importance of having available a theoretical framework that can be used to interpret an empirical finding and to integrate it with findings from other parts of economics. Development economists have for many years felt the need to go beyond the neoclassical model to understand the experience of developing countries. They have typically done so using models that are not consistent with the theory and evidence developed in other parts of economics. For example, they have assumed a Harrod-Domar technology with fixed coefficients and an excess supply of labor, or by used disequilibrium models in which wages and prices are not equated across sectors. The difficulty with this approach is that there is not enough data in development proper for it to operate as a closed system that does not take advantage of the findings from the rest of economics. The model proposed here at least hints at the possibility of integrating the analysis of growth and development across all countries and across the disciplines of industrial organization, firm level studies on research and development, and possibly even with the recent work in macroeconomics on aggregate production and departures from price taking (e.g. Hall, 1989).

References

Arrow, K. J. 1962a. "The economic implications of learning by doing." **Review of Economic Studies**, 29 (June), 155–73.

——————. 1962b. "Economic welfare and the allocation of resources for invention." In **The Rate and Direction of Inventive Activity**. Princeton: NBER and Princeton University Press.

Azariadis, Costas and Allan Drazen. 1988. "Threshold externalities in economic development." University of Pennsylvania Working paper, October 1988.

Barro, R. J. 1989a. "A cross country study of growth, savings, and government." NBER working paper 2855.

——————. 1989b. "Economic growth in a cross section of countries." Harvard University working paper.

Becker, G. and K. Murphy. 1988. "Economic growth, human capital, and population growth." University of Chicago, June 1988.

Chamberlain, G. and Z. Griliches. 1974. "Unobservable with a variance-components structure: Ability, schooling, and the economic success of brothers." **International Economic Review**, 41(2), 422–9.

——————————. "More on brothers." In P. Taubman ed. **Kinometrics: The Determinants of Socio-Economic Success Within and Between Families**. Amsterdam: North Holland, 97–124.

Dixit, Avinash K. and Joseph Stiglitz. "Monopolistic Competition and Optimum Product Diversity." **American Economic Review** 67 (June), 297–308.

Ethier, W. J. 1982. "National and International Returns to Scale in the Modern Theory of Trade." **American Economic Review** 72 (June), 389–504.

Jamison, D. and L. Lau. 1982. **Farmer Education and Farm Efficiency**. Baltimore: Johns Hopkins University Press.

Hall, Robert E. 1989. "Invariance Properties of Solow's Productivity Residual." NBER working paper No. 3034.

Lucas, R. 1988. "On the mechanics of economic growth." **Journal of Monetary Economics** 22, 3–42.

Romer. P. 1986. "Increasing returns and long run growth." **Journal of Political Economy** 94 (October), p.1102–1037.

——————. 1987. "Crazy explanations for the productivity slowdown." S. Fischer ed. **NBER Macroeconomics Annual**, Cambridge MA: MIT PRess.

——————. 1988. "Endogenous technological change." University of Chicago working paper.

————. 1989a. "Capital accumulation in the theory of long run growth." In R. Barro ed. **Modern Macroeconomics**. Cambridge: Harvard University Press.

————. 1989b. "Measurement error in cross-country data." Working paper.

Schumpeter, Joseph A. 1942. **Capitalism, Socialism and Democracy**. New York: Harper and Brothers.

Shell, Karl. 1967. "A model of inventive activity and capital accumulation." In K. Shell, ed. **Essays in the Theory of Optimal Economic Growth**. Cambridge MA: MIT Press.

Solow, Robert M. 1956. "A contribution to the theory of economic growth." **Quarterly Journal of Economics** 70 (February), p. 65-94.

Summers, Robert and Alan Heston. 1988. "A new set of international comparisons of real product and price levels: Estimates for 130 countries, 1950 to 1985." **Review of Income and Wealth** (March), 1-25.

Wilson, R. 1975. "Informational economies of scale." **Bell Journal of Economics** 6 (Spring), 184-195.

Table 1: Variable Definitions

C	A constant term used in all of the regressions.
Y60	Real per capita income in 1960, measured in 1980 dollars. Constructed using current price weights. From Summers and Heston, RGDP2. Range, \$250 to \$7400.
GROWTH	The average annual rate of growth of Y60, measured in percentage points times 100, over the years 1960 to 1985. Range, -4 to 7.
GOV	Share of GDP devoted to government spending on items other than investment goods, in percentage points times 100, averaged over the years 1960 to 1985. Measured as the ratio of current price government spending to current price GDP. Range, 5 to 35.
INV	Share of GDP devoted to investment, averaged over 1960 to 1985. Measured as the ratio of current price investment to current price GDP. Range, 4 to 37.
LT60	Percentage of the population times 100 that is literate in a survey year close to 1960. Range, 1 to 98. From UNESCO.
LT_DIFF	Change in the literacy rate between 1960 and 1980, in percentage points times 100. Range, -19 to 56. From UNESCO
NP60	The logarithm of consumption of newsprint per capita in 1960. Range, -4 to 4. From UNESCO.
RD60	The number of radios per 1000 inhabitants in 1960. From UNESCO.

Table 2

```
Least Squares:  Dependent Variable is GROWTH

Number of observations: 94
=====================================================================
     VARIABLE     COEFFICIENT    STD. ERROR       T-STAT.   2-TAIL SIG.
=====================================================================
        C          1.8860644     0.6677856       2.8243562     0.006
       Y60        -0.0006255     0.0001744      -3.5853773     0.001
       INV         0.1883967     0.0274488       6.8635803     0.000
       GOV        -0.1165580     0.0299402      -3.8930302     0.000
      AFDUM       -0.8970903     0.4440245      -2.0203621     0.047
      LADUM       -1.2957244     0.4021405      -3.2220689     0.002
       LT60        0.0154671     0.0081786       1.8911618     0.063
=====================================================================
R-squared                 0.583406    Mean of dependent var    1.811394
Adjusted R-squared        0.554675    S.D. of dependent var    2.063386
S.E. of regression        1.376952    Sum of squared resid     164.9518
Durbin-Watson stat        2.520105    F-statistic              20.30605
=====================================================================
```

Table 3

```
Instrumental Variables:  Dependent Variable is GROWTH

Number of observations: 69
Instrument list: C INV GOV AFDUM LADUM RD60 NP60
=============================================================
    VARIABLE    COEFFICIENT   STD. ERROR      T-STAT.   2-TAIL SIG.
=============================================================
        C        2.0456095    0.9568901    2.1377686      0.037
       Y60      -0.0002062    0.0002799   -0.7367500      0.464
       INV       0.1475381    0.0352431    4.1863005      0.000
       GOV      -0.0926471    0.0355311   -2.6074894      0.011
      AFDUM     -1.2276418    0.6052440   -2.0283420      0.047
      LADUM     -1.3738536    0.4143630   -3.3155794      0.002
      LT60       0.0061744    0.0172710    0.3575018      0.722
=============================================================
```

Table 4

```
Instrumental Variables:  Dependent Variable is GROWTH

Number of observations: 69
Instrument list: C GOV AFDUM LADUM RD60 NP60
======================================================================
     VARIABLE    COEFFICIENT    STD. ERROR       T-STAT.   2-TAIL SIG.
======================================================================
        C          2.2644223     1.1360664      1.9932130      0.051
       Y60        -0.0004166     0.0003321     -1.2545811      0.214
       GOV        -0.0508671     0.0405552     -1.2542699      0.215
      AFDUM       -1.0275136     0.7179690     -1.4311393      0.157
      LADUM       -1.8331959     0.4769421     -3.8436447      0.000
       LT60        0.0386495     0.0187536      2.0609111      0.044
======================================================================
```

Table 5

```
Instrumental Variables: Dependent Variable is INV

Number of observations: 51
Instrument list: C GOV AFDUM LADUM LT_DIFF NP60 RD60
==============================================================
   VARIABLE    COEFFICIENT   STD. ERROR      T-STAT.   2-TAIL SIG.
==============================================================
      C         1.3745661    4.3981236     0.3125347     0.756
     Y60       -0.0001662    0.0008864    -0.1874816     0.852
     GOV        0.1684469    0.1368366     1.2310076     0.225
    AFDUM       1.0307406    2.8535393     0.3612148     0.720
    LADUM      -1.6100160    1.4534961    -1.1076851     0.274
    LT60        0.1574893    0.0610410     2.5800592     0.013
   LT_DIFF      0.2044889    0.0879624     2.3247317     0.025
==============================================================
```

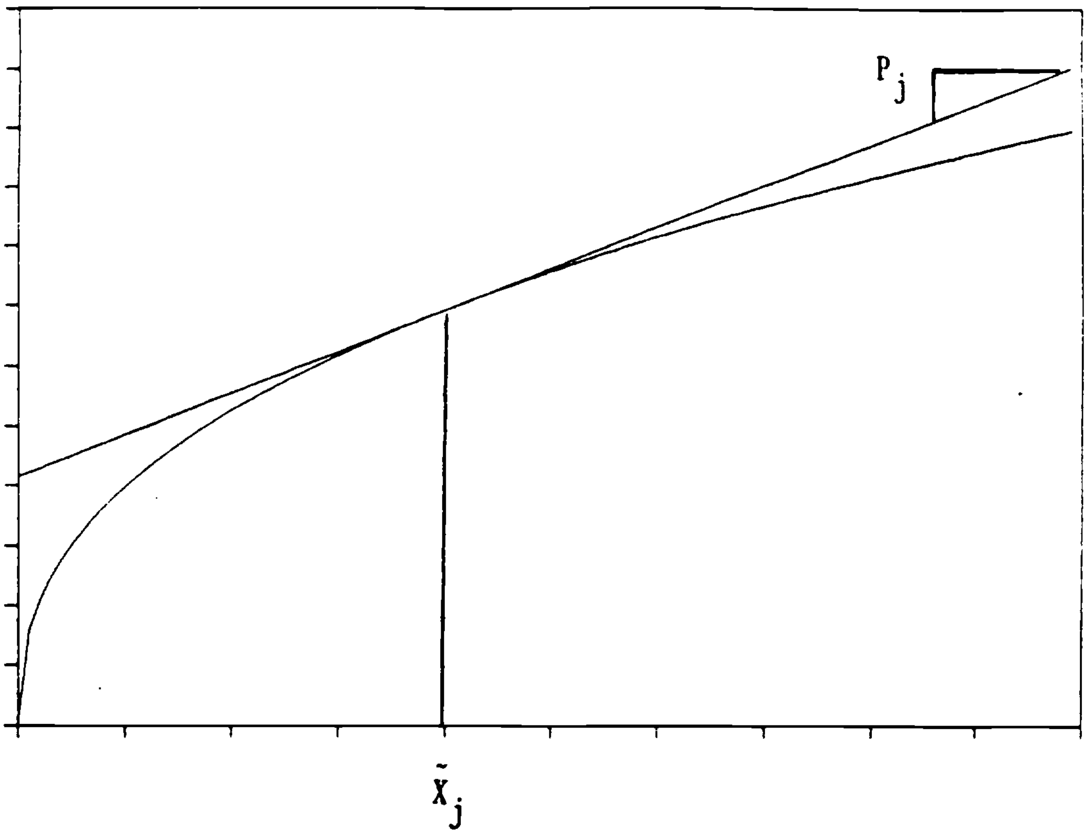

Figure 1