PN ABD-361
62856

Human Capital and Growth: Theory and Evidence

Paul M. Romer

University of Chicago

April 1989

Prepared for the April 1989 Carnegie-Rochester Conference. This work was supported by NSF Grant SES 8821943 and by a Sloan Foundation Fellowship.

This paper offers theory and evidence on the connection between human capital variables and cross country variation in growth rates. Section 2 below presents the outline of a framework that organizes the subsequent discussion. It conclusions can be simply stated. In a model that 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 rates of change of inputs and outputs suggested by growth accounting, there will be a role for the level of human capital variables in explaining the rate of growth of output and the rate of investment. In a regression equation that tries to estimate separate roles for both investment and human capital variables in explaining the rate of growth, collinearity may cause the human capital variables not to enter in the equation They should still have explanatory power for investment.

The empirical part of the paper (Section 3) focuses exclusively on the implication that the level of a human capital variable like literacy has a distinct explanatory role in cross country regressions for per capita income growth. The theoretical section serves only to motivate this hypothesis, and empirical section can be read independently. Tests of the implications for investment are postponed for later inquiry. The conclusions from this analysis can be summarized as follows:

1) There are results that can be interpreted to mean that the initial level of literacy and its rate of growth are positively related to per capita income growth. However, these results can more plausibly be interpreted to mean that there is substantial mismeasurement in the estimates of the level of income across countries that biases the

estimates. Attempts to estimate an effect for literacy that are not subject to the problem of measurement error in the level income seem to face a serious problem with multicollinearity.

2) The rate of investment has a robust positive association with the rate of growth. Under the interpretation that takes the results for human capital at face value, its magnitude is on the order one would expect from a standard growth accounting model if investment is exogenous. Under the alternative interpretation, its coefficient is about twice as high as a growth accounting calculation would suggest.

Other substantive implications that have a bearing on related models and empirical work are:

3) The level of government spending on items other than investment seems to be negatively related to the rate of growth, but the estimated magnitude depends very much on which interpretation one adopts of about the problem of measurement error. Under one interpretation, the effect of government is very large, and very sensitive to the use of an estimator which corrects for the fact that for some countries, government spending can grow through direct international transfers that are not associated with domestic tax increases.

4) Because of the possibility of measurement error in the level of per capita GDP in the early years of the sample, it is difficult to draw firm conclusions about the effect of the level GDP on the rate of growth. In particular, there is no unambiguous evidence that low income countries

tend to catch up with high income countries when other variables like investment are held constant.

5) Dummy variables for Africa and Latin America that have been found to be significant in some previous specifications are not always significant here, especially if one makes allowance for the possibility of measurement error in the initial level of per capita income. The finding of a negative dummy variable for Latin America remains a puzzling and relatively robust finding.

The methodological conclusions, which are perhaps the most robust findings here, include:

6) Errors in variables may be very important in cross country analyses. For many of the variables of interest, there are other variables that can be used as instruments. In several important cases, an instrumental variables estimate is quite different from the least squares estimate.

7) In a regression of growth rates on other variables, there is evidence of heteroskedasticity that is related to the indicators of data quality provided by Summers and Heston. There is some evidence that possible errors in the estimates initial level of in per capital GDP and of the share of government in GDP are related to the indicators of data quality, but this is not the only interpretation of this evidence.

8) Finally, for the analysis here spanning 25 years of data, it makes an important difference whether one uses data on the share of government and

investment in GDP that are measured using current price weights or using fixed price weights from a particular year.

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 for 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, Romer (1986) focuses attention on a mongrel notion of aggregate capital that combines elements of both knowledge and physical capital, and that 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) focuses 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.

A dogmatic adherent logical positivism would object that these models are not operational, that they do not specify a cookbook list of instructions together with a predicted outcome that could be used to test the model. Therefore, the positivist would argue, they are akin to metaphysics and have no scientific content. This judgment may be too harsh. An examination of how science actually works shows that the positivist viewpoint misses much of the richness of the interaction between theory and evidence, and it largely discredited (except mysteriously, among economists.) But without going to the extremes of the logical positivists, it is easy to be sympathetic with the view that models that lend themselves more readily to the analysis of available data would be welcome.

This section outlines an attempt at such a model. It builds on the model outlined in Romer (1988), and extends its 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, it is possible to make informed conjectures about how the extended model will behave, but none of these conjectures are verified rigorously here. For the most part, what this kind of extension can do is detail a list of possible variables to use and possible interactions to look for in the analysis of data. Even in its very sketchy form, the model outlined here serves a purpose, for it suggests specifications of equations that many not at first seem obvious and that are not suggested by the conventional growth accounting framework. In particular, it forces one to move beyond a narrow focus on the rates of change of inputs, and suggests that the levels of some inputs may be related to rates of growth.

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 bears on them. To keep the scope manageable, the model and the subsequent empirical analysis will neglect 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 (1989). Once the issues considered here are better understood, it should be possible to consider an extension that includes the model here and the models considered by others as special cases.

2.2 The Model

Let M denote the number of individuals in a closed economy, and let i denote a typical individual. Each individual has a fixed allotment of time in any given period that can be divided between two different kinds of educational activities, and four different productive activities. Leisure is of course possible as well, but this will not be explicitly noted. Every individual 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 of course it could be more explicitly modeled as the outcome of investments in nutrition, health care, and other inputs. E_i for each individual will be measured as it is in the data, in total years of schooling. Thus, for the individual, E_i grows according to

$$\dot{E}_i = \begin{cases} u^E & \text{if } E_i \le 12, \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

where $u^E \in [0,1]$ denotes the fraction of an individuals 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. In any empirical application, variables will of course be measured over discrete intervals.) If the average level of education in the population is denoted as

$$E = \frac{1}{M} \sum_{i=1}^{M} E_i, \qquad (2)$$

the rate of growth of E in the population as a whole will be

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

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

By convention, scientific skills S_i are distinguished from skills acquired from primary and 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 suggest 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:

$$\dot{S}_i = \begin{cases} u^S & \text{if } E_i = 12, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$\bar{S} = \frac{1}{M} \sum_{i=1}^{M} S_i, \tag{5}$$

$$\dot{S} = \sum_{i=1}^{M} \dot{S}_i - \delta M\bar{S}. \tag{6}$$

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. The key feature of this specification is that both of the variables E and S are bounded on a per capita basis. In particular, neither can exceed the average length of life of the individuals in this economy. For unbounded per capita income growth to take place, some input will have to grow without bound on a per capita basis. Average years of primary, secondary or postgraduate schooling are not candidates for this kind of variable. The fact that they cannot grow forever should not obscure the fact that in actual data they may exhibit important growth in the relevant sample period.

Total output of potential consumption goods in this economy will be denoted as $Y \in \mathbb{R}$, and expressed as a function of labor inputs $L^Y = \Sigma_i u_i^Y L_i$, educational inputs $E^Y = \Sigma_i u_i^Y E_i$ and a list of intermediate inputs

$X = (X_1, X_2, \ldots)$. As usual, Y must be split between consumption and investment. Since u_i^Y denotes the fraction of time individual i devotes to production of Y, this individual must supply all three of $u_i^Y L_i$, $u_i^Y E_i$, and $u_i^Y S_i$ to this sector. By assumption, scientific skills make no contribution to increased output of Y, 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 in Becker and Murphy (1988), and is not pursued here.

Let Y denote output net of the amount of investment needed to maintain the capital stock, so that

$$Y(L^Y, E^Y, X^Y) = \dot{K} + C. \tag{7}$$

Because all of the other goods specified in the model are intermediate inputs into production of Y, Y is like a measure of net national product. (If K were the only durable productive input in the model, this would be identical to net national product, but durable intangible inputs will be introduced below, and their rates of accumulation should enter into net national product as well.)

Typically, one would let capital, measured as cumulative foregone consumption, enter directly as an argument into the production function for Y. In the specification used here, capital enters indirectly through the list of intermediate inputs X. A typical component of this list X_j could refer to lathes, computers, or trucks. It simplifies the accounting to let X_j

stand for the flow of services from lathes, computers or trucks available at a point in time, so that X_j is not itself a durable even though durable capital is used to produce it.

To allow for the fact that new intermediate inputs can be introduced as growth takes place, the list X of actual and potential inputs is assumed to be of infinite length. At any point in time, only a finite number of X_j's will be produced and used in positive quantities. For example, if X_J denotes the services of a DOS based personal computer with a 40 Mhz clock speed, X_J is equal to 0 as this is written because no such computers are available (yet). One can nevertheless makes conjectures about how its availability would affect output if it were. The assumption that the function Y and the complete infinite list of arguments X_j is known with certainty is of course not to be taken literally, but it is likely that the main points of the analysis that follows will carry over to a model with uncertainty about these elements.

For a particular intermediate input of type j that is already in available, the flow of output X_j can be written as a function of the amount of capital K^{X_j}, physical labor $L^{X_j} = \Sigma_i u_i^{X_j} L_i$, and education skills $E^{X_j} = \Sigma_i u_i^{X_j} E_i$ that are employed. Scientific skills are assumed not to enter into any of the manufacturing processes for Y or for the X_j's.

There is probably little harm in assuming that the production functions $Y(\cdot)$ and $X_j(\cdot)$ are homogeneous of degree 1. 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 departures from the usual assumptions about returns to scale seem inevitable is in the process whereby new goods are produced. The essential observation here is that the introduction of a new good involves expenditures that are quasi-fixed. They must be incurred to produce any goods at all, but they do not vary with the level of production. Generically, these costs can be thought of as designs, mechanical drawings, or blueprints. The manufacturing function $X_j(\cdot)$ then describes what happens when these drawings are sent down to the machine shop or factory floor for production.

The distinction drawn between rivalry and excludability in the study of public goods is very useful in this context. The key feature of something like a design is that it is a nonrival input in production. That is, the use of a design in the manufacturing of one lathe, computer or truck in no way limits or interferes with its use in the production of another lathe, computer, or truck. The extent of rivalry is something that 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.

Despite periodic acknowledgements that nonrivalry is inherent in the idea of knowledge or technological change (e.g. Arrow 1962, Shell 1967, Wilson 1975), models of growth have tended to neglected this issue. The original Solow (1959) model of exogenous technological change implicitly acknowledge the nonrival aspects of knowledge, but did so ruling out the possibility that it was privately provided. Arrow (1962b) allows for nonrival knowledge, but relies on a learning by doing formulation the makes knowledge privately provided, but only by accident. Romer (1986) and Lucas (1988) introduce kinds of knowledge that are partly excludable and rival, and partly nonexcludable and nonrival. Once again, nonrival knowledge is produced only as a side effect of some other activity. These attempts to finesse the issue of the private provision of nonrival inputs presumably arise from the technical difficulty that nonrival goods, especially privately provided nonrival goods, present for economic models rather than a conviction that 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, and almost all of the output from this activity has the nonrival character of blueprints, designs or inventions.

A casual examination of the business press suggests that the problems for individual firms created by the private provision of a nonrival input are very real. In the last month there have been stories about thefts of secret process technologies used by Du Pont in the production of Lycra, and of thefts of box loads of documents from Intel concerning its 80386, 80387 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 an extensive criminal and civil proceedings to keep its $200 million dollar 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, p.1, August 16, 1988.)

The nonrivalrous aspect of new good design is captured here by assuming that there is an additional variable A, representing the outcome of applied research and development, which measures the stock of designs. ("A" for applied.) A fixed increment to A, the design for good j, must be produced before it is possible to start production of X_j. Once it is acquired, the level of production X_j depends only on its direct inputs $X_j(L^{X_j}, E^{X_j}, K^{X_j})$.

The production technology for the designs or blueprints captured in the good A is assumed to depend on the amount of scientific and educated labor S^A and E^A used in this process, together with the list of intermediate inputs X^A used for this purpose, the existing stock of A and the stock of an additional nonrival input B (for basic). The stock B is intended to capture the basic research that is exploited in applications. Its production depends on the amount of scientific talent S^B devoted to this activity, its own level B, the level of the applied stock of knowledge A, and any of the intermediate inputs X that are available for use. Thus,

$$\dot{A} = \dot{A}(E^A, S^A, A^A, B^A, X^A), \tag{8}$$

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

In both of these functions, the intermediate inputs may not have the same productivity as they have in producing Y, or any productivity at all. Computers matter for the production of A and B; turbine blades do not.

There is a further extension that is not pursued in detail here. To model learning by doing, arguments in the production of Y or of 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 plus the educational level E^Y would appear as arguments of $\dot{A}$. This extension would tend to reinforce the conclusion highlighted below that an increase in the level of education in the economy as a whole may tend to increase the rate of growth. For many policy questions, it is important to establish the relative importance of direct investment in $\dot{A}$ versus indirect learning by doing investment, but for the empirical work undertaken below, all that matters is that learning by doing will add an additional channel through which the level of E can affect growth.

The constraints on the rival inputs in this model are straightforward. At the individual level, the constraint on the allocation of time is

$$u^Y + \Sigma_j u^{X_j} + u^E + u^S + u^A + u^B \leq 1. \tag{10a}$$

If L, E, S, and X denote the total stocks of the rival goods, the aggregate adding up constraints are:

$$\begin{aligned} L^Y + L^X &\leq L, \\ E^Y + E^X + E^A &\leq E, \\ S^A + S^B &\leq S, \\ X_j^Y + X_j^A + X_j^B &\leq X_j \quad \text{for all } j. \end{aligned} \tag{10b}$$

The constraints on the nonrivalrous goods are of course different:

$$\begin{aligned} A^A \leq A,\ A^B \leq A, \\ B^A \leq B,\ B^B \leq B. \end{aligned} \tag{10c}$$

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.

It should be clear that 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 individual inputs corrected for double counting (i.e. for the production of the same piece of A or B by different firms or labs).

At the level of generality used here, there is not much that one can prove rigorously about this system of equations. However, one immediate implication of the presence of nonrival inputs in production is that the competitive assumptions needed for a complete accounting for growth do not hold. At the firm level, this shows up in decreasing average costs of producing X_j that arise because of the initial fixed investment in design costs. If the firm priced output at marginal 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, 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, 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 Y and all of the X_j's, and replicates the accumulation of E and of S that takes place in the first economy. Since the underlying production functions for Y 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 available in the original economy. Even if that portion of the talent E and S that is 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 and S resources were instead used to produce additional units of A or B or merely used in production of Y 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, S, and K alone. If one allows for simultaneous increases in A and B as well, the argument for increasing returns is that much stronger.

The fact that it is not possible to replicate you, me, or any number of other existing resources is not relevant here. All that matters from this thought experiment is what it can reveal about the underlying mathematical properties of production. What it shows is that it is not possible for market prices to 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.

This result carries over into this more complicated dynamic setting. If V(L,E,S,K,A,B) denotes the present value of a Pareto optimal stream of consumption starting from given stocks of inputs, then

$$\frac{\partial V}{\partial L}L + \frac{\partial V}{\partial E}E + \frac{\partial V}{\partial S}S + \frac{\partial V}{\partial K}K + \frac{\partial V}{\partial A}A + \frac{\partial V}{\partial B}B > V(L,E,S,K,A,B).$$

The price of each asset, or equivalently, the present discounted value of the stream of earning from the different types of human capital, cannot be equal to the marginal social product of this good. This has the positive implication that growth accounting exercises that equate marginal values with prices will fail. It has the normative implication that except in the very unlikely case that L is the only factor that is undercompensated, the accumulation of some or all of the other factors will most likely take place at a rate that is too low.

So far, the discussion of the model has been vague about the form of equilibrium that obtains and about where the increases in A and B come from. The easiest case to consider, and one that illustrates clearly the claim made above, is one where both A and B are nonexcludable, and hence cannot be privately provided. In the more usual (but less explicit) terminology, they are said to have purely external or pure spillover effects. Suppose further that increases, if any, in A and B arise from government revenue collected through lump sum taxes. Then $\dot{A}$ and $\dot{B}$ will be functions of the path of funding chosen by the government, and could potentially be exogenously determined relative to other economic variables in the system, in which case the model looks very much like one with exogenous technological change. In this case it is relatively easy to see why growth accounting must leave an unexplained residual whenever A grows.

Figure 1 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. If the price of this input is P_J and the firms producing output are price takers, X_J will be used at a level $\tilde{X}_J$ such that its marginal productivity is equal to P_J. If X_J is increased by a small amount ΔX_J, its effect on total output can be approximated by $P_J \Delta X_J$. If the producers of X_J are also price takers so that P_J is the marginal cost of X_J, the increase $P_J \Delta X_J$ is equal to the value of the additional inputs L, E, K needed to produce the increase ΔX_J. If there is a increase in the aggregate stocks of the inputs L, E, K in the absence of any change in A, it would be spread over increases in all of the existing inputs. The effect on total output would be the sum across all of the different inputs of these kinds of effects, with the net result that the change in total output would be approximately equal to the value of the increase in the initial inputs. Thus, current prices times the increase in the quantities L, E, K that are used give a good approximation to the increase in aggregate output. Thus, if there is no government funding and A stays constant, growth accounting will not leave any residual.

Now, consider what happens if the government supplies a design for a new good J to the market. Let X_J increase from 0 to the level X_J, and suppose that some large fraction of all of the increase in the inputs L, E, and K in a given period was devoted to producing the new intermediate input. Under marginal cost pricing of X_J, the value of the increase in the inputs L, E, K used to produce X_J will still be equal to P_J times the increase in X_J. But in this case, the large change in the quantity X_J from 0 to $\bar{X}_J$ means that $P_J \Delta X_J = P_J \bar{X}_J$ is not a good estimate of the resulting increase in output. The first unit of X_J has a marginal effect on output that is much larger than P_J. As the figure shows, the increase in Y is the vertical distance ΔY, which is substantially larger than the value $P_J \bar{X}_J$. Any growth accounting exercise will underestimate the growth in output. If increases in A take place every period, growth accounting would find a residual in the sense that the rate of growth of output would be persistently higher than the rate of growth of the value of inputs. Moreover, the magnitude of this unexplained residual will be increasing in the rate of growth of A.

(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. However, since welfare is not measured, this effect has no obvious implications for the analysis of cross country data on growth.)

The accounting described above does not take any account of the resources that the government uses to produce the increases in A each period, but it is clear that additional A can be produced in each period holding constant the inputs used for this purpose. A fixed stock of scientific and educated talent could presumably continue to produce increases in A and B indefinitely. By this logic, the rate of increase in A will be an increasing function of the level of inputs used in $\dot{A}$ and $\dot{B}$. This is the new relationship alluded to above, one that has no counterpart in growth accounting. The unexplained component of the rate of growth will be a function of the level of the stocks of resources devoted to research and development. In addition, the rate of investment in new K should be positively related to the rate $\dot{A}$ at which new opportunities for investment are introduced. Thus, $\dot{A}$ affects not only the residual from growth accounting, but also the rate of increase of the input K. One would ideally try to relate the rate of growth of a variable like A to the rate of growth of output and of K. In the absence of internationally comparable data on new good introductions, innovations, or patents, one could still compare the level of government support for science with the rate of growth of output and of capital.

As a cross country model of growth this model is surely wrong on two important counts. First, it flies in the face of evidence that the vast majority of expenditure on A is privately financed. Second, it neglects the fact that countries are not closed economies that operate in isolation. This implies at least that the stock of B that is relevant for a given country should be the entire worldwide stock, not just the locally produced stock. (It also means that the extent of integration with world markets is an important determinant of income and growth as noted in Romer 1988 and explored in Grossman and Helpman 1988, but the interaction between trade and growth is another of the connections that cannot be pursued here.) For almost all of the countries in the sample considered below, it is sufficient to treat the rate of growth of B as exogenously given, determined in a small number of very rich countries.

That said, there is still every reason to believe that the process of producing A, of designing specific goods that can be sold and processes for manufacturing these goods, is very important for all of the countries in this sample. If the results of basic research had direct value in production, the assumption that the is would reduce the model to one with exogenous technological change for most countries, but the mere fact that a country can subscribe to all of the scientific and engineering journals in the world does not ensure that growth can take place if there is no local educated and scientific talent to convert this basic knowledge into a form that leads to the production of new goods in a particular economic environment. What is used in production is applied designs A. B is an input in the production of A, but it is not the only one.

Casual evidence suggests that in almost all cases, the local production of A inherent in the adaptation of technology to the production of new goods it undertaken by private firms, not by governments. Thus, in the absence of direct evidence on the rate of production of new goods, one should not expect to find that government expenditure on support for science and engineering is an important variable for explaining cross country variation in growth of output or of capital. In many case there is essentially none at all. However, one can argue that the total stock of educated and scientific talent in a country should be related to the quantity allocated to the production of A and therefore to growth in output and capital.

This result can be explicitly derived in a different special case of the general model used here. Romer (1988) assumes that something like A is excludable (at least as it applies to the production of X_j's) and therefore is privately financed. The specific model combines the variables E and S into a single human capital variable H and assumes that its level is constant. It also combines basic research B and applied product development A into a single variable A. A very simple specification of the functional forms for $Y(\cdot)$, $X_j(\cdot)$, $\dot{A}(\cdot)$ is used, one that relies heavily on an artificial symmetry between all the goods X_j. This results in a simple form of strategic interaction between the different firms that are the unique suppliers of the goods X_j. The result is an industry equilibrium with a familiar form of monopolistic competition. Producers of new goods can recoup their initial design costs by charging a price for their unique good that is higher than marginal cost.

This institutional setting shows how it is that private production of a nonrival good like A can take place. Because it is simple, it also permits explicit derivation of the determinants of rate of growth. In this special case, increases in the total stock of trained human capital lead to increases in the amount of human capital 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.

Having an explicit solution in this special case also gives a warning about the interpretation of empirical results of the model. 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 A. Thus, in a regression that relates the rate of growth of output to the rate of growth of K and to the level of education and scientific talent, collinearity between $\dot{K}$ and $\dot{A}$ will mean that there is nothing left for the level of education and scientific talent to explain. $\dot{K}$ will have a coefficient that is bigger than a growth accounting model would predict because it picks up both the direct effects of increases in K and the effects of increases in A.

In more general models, it need not be the case that $\dot{K}$ and $\dot{A}$ are perfectly collinear, so a separate effect for E and S could be observed. In any case, the model has the additional implication that the rate of growth of $\dot{K}$ should be explained in part by the level of E and of S.

In summary, the novel empirical implications of this analysis are that both the rate of growth of per capita income and the rate of investment will be positively related to the level of human capital variables like education or scientific talent. It is possible that the schooling variable will not be significant in a regression that also includes the rate of investment. If so, the rate of investment should have an apparent effect on output that is large compared to the one implied by the share of capital in total income. For the usual growth accounting reasons, one might also expect that the rates of change of these variables will be positively related to growth, but this is not certain. Because S is assumed only to affect $\dot{A}(\cdot)$ and $\dot{B}(\cdot)$, growth in S will not have any effect on Y once changes in A are accounted for. To the extent that E does not appear in $Y(\cdot)$ or $X_j(\cdot)$, and only appears in $\dot{A}(\cdot)$, growth in E will not have a large independent effect on Y either.

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 direct measures like literacy and indirect measures like life expectancy and per capita consumption of newsprint. 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. In fact, even the analysis of the effects of literacy on investment are deferred, although preliminary results are described briefly in the last section. Thus, the current results are concerned only the connection between basic literacy and the rate of growth of per capita income. As will become clear, this very narrow focus is dictated by the difficult issues of interpretation that arise in this context. In any extension, the exposition of regression results risks becoming impossibly long. As it is, the paper reports more regression results than any one person (including the author) can keep track of in his or her head.

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 in Romer (1987, 1989). These data have also been used in conjunction with detailed data on government expenditure and demographic variables by Barro (1989) in an analysis that focuses on fertility choice and on a possible productive role for government investment expenditure. In what follows, some comparisons with results from Barro will be drawn, but it should be understood that none of these results are strictly comparable. His estimates make use of variables that are not used here. Also, because of the limited of data availability for some variables, the sample of countries considered is not generally the same. This problem recurs throughout all of the subsequent analysis. Anytime an additional variable other than one from the Summers and Heston data set is used, the number of countries with complete data gets smaller.

Other than Barro, the work most closely related to the results reported here is work of Hicks (1979,1980) and the preliminary regressions reported in Azariadis and Drazen (1988.) To the extent that they produce comparable results, the regressions reported below generally reproduce their findings, but additional evidence reported here calls the interpretation they offer into question.

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 this 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, some impressions can be offered. 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 (as it is now practiced) 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 signaling is not an important issue, and inputs and outputs can be measured relatively directly. This leaves open the possibility that unmeasured individual attributes cause both the variation in educational achievement across individuals and the variation in productivity, but there is separate evidence like that in Chamberlain and Grilliches (1974, 1979) using sibling data on education, labor market outcomes, and test scores that suggests 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, these seems to be a general sense that the "human capital revolution" in development has been something of a disappointment, and that growth accounting measures of the effects of education do not help us understand much of the variation in growth rates and levels of income observed in the world. One illustrative finding is that of Barro (1988) that school enrollment rates were were closely negatively related to the rate of growth of the population, neither these enrollment rates nor the ratio of government expenditure on education to GDP had any explanatory power in the regressions for per capita income growth rates. 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, it must be admitted that the results reported in this first step will not by themselves redeem education, but as noted at the end, preliminary evidence about the effects of education on investment appear to be more promising.

3.2 Regression Results

The list of variables used in the subsequent regressions is given in Table 1. 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. However, much of the subsequent analysis turns of the properties of the initial level of per capita real income in 1960, and at roughly $50,000 (in 1980 dollars) Kuwait is an outlier by an order of magnitude. The next highest value if 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. Rather than let Kuwait 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. Table 2 lists them, together with a measure of data quality provided by Summers and Heston.

The basic starting point for the analysis is the regression described in Table 3. Several remarks are in order before turning to this table. It gives two stage least squares (equivalently, instrumental variables) estimates of the effects that the average share of total investment (including government investment) in GDP over the sample period, the average share of noninvestment government spending as a share of GDP, and the level of literacy in 1960 have on per capita income growth from 1960 to 1985. The regression includes several nuisance parameters for which there is little theoretical support, but which have important interactions with the variables of interest. Following the lead of Barro, the initial level of per capita income is allowed to influence growth in an arbitrary way. This is accomplished by letting the level of income in 1960 (RY260), this level squared (RY26Q, "Q" for quadratic), and the log of this level (RY26OL, "L" for logarithm) all enter in the equation. Since Barro found that dummy variables for the continents of Africa and Latin America (including Central America and Mexico) had significantly negative effects on growth, they are included here as well.

It is not clear how to interpret the coefficients of these variables, and it will become even less clear as more evidence is presented. However, one useful way to interpret the coefficients on the other variables is to recall that in a multiple regression of a variable Y on two sets of variables X_1 and X_2, the coefficient on X_2 can be estimated by regressing both Y and X_1 first on X_2, then regressing the residuals from this step on each other. Thus, the coefficient on, say, the share of investment is exactly what one would calculate if the share of investment was replaced by deviations of investment from the share that would be predicted from a regression on the initial level of income, its square, and its logarithm.

This interpretation explains the motivation for allowing a very flexible dependence of the variables on RY260. Because the three forms of RY260 are closely correlated, the individual coefficients are not precisely estimated, but they are jointly highly significant. Excluding one or two of these variables did not affect any of the other inferences.

The use of instrumental variables estimators was motivated by a concern that measurement errors could be a serious problem in these data, and by the observation that many of the variables of interest had associated with them variables that provide at least partially independent measurements of the underlying concept of interest. For example, all of the series from Summers and Heston come in a form that is calculated using 1980 prices weights for the different components of GDP (RY160, CONS, INV, GOV in the notation of this paper or RGDP1(1960), c, i, g in the notation used by Summers and Heston) and a form that is measured using current prices (RY260, CCONS, CINV, CGOV in the notation used here or RYGDP2(1960), cc, ci, cg in the notation of Summers and Heston.) Following their lead, the prefix "C" is used here to indicate that current price weights were used.

The analysis here proceeds under the assumption that the quantities valued in current prices are better indicators of the underlying quantities of interest, but allows for the possibility that each of the possible measures is contaminated with some error associated with index number problems caused by changing relative prices. (Note that the use of these kinds of instruments will not correct for any measurement errors in the basic data that are common to the two measures provided by Heston and Summers. This issue is considered further below.) To the limited extent that this issue was explored, the use of the 1980 price weight data as instruments for the current price data did not have a large effect on any of the inferences, but neither did they cause much loss in efficiency, so they were maintained throughout.

It is important to note that the difference between current prices and fixed year prices are not in every case a trivial matter. When the basic data were extracted from the Heston Summers data set, the following result was noted for the first country in the table, Algeria. If one averages the share of government, consumption and investment over the period 1960 to 1985, the current value measures indicate that on average, the share of net exports in GDP was equal to -17%. Using the measures that are based on 1980 price weights suggested that Algeria had on average net exports that were positive and equal to 3% of GDP. Evidence that the current prices may be better is offered below in Table 5, so throughout the rest of the analysis, current price variables are used in as the basic variables, and 1980 price variables are used as instruments.

The other variable in the regression in table 3 that is associated with an instrument is the initial level of literacy. The concern here was that literacy might not be measured in strictly comparable ways across different countries, and that the reported measures would therefore contain measurement errors relative to the true measure of interest. The two instruments that were thought to offer an independent indication of the level of effective literacy in a country are the level of life expectancy and the per capita consumption of newsprint. Because the distribution of values for per capita consumption of newsprint turns out to be very significantly skewed, the logarithm of the per capita level NP60L ("NP" for newsprint, "60" for 1960, "L" for logarithm) was actually used as the instrument in the equations.

Initial experimentation confirmed that the logarithm performed better as an instrument than did the level. (Experimentation with the log of literacy versus its level revealed that the level provided a slightly better fit to the data, but the difference is not large.) Table 3 reports results using life expectancy as the instrument rather than the newsprint variable. The results in the two cases were generally similar, and an indication of the differences is given in the subsequent discussion.

In principle, one could use both variables as instruments for the level of literacy, but because the coverage of the two variables is incomplete and not identical, the use of both results in the exclusion of additional observations. In every regression, any country which did not have complete data on one of the variables under consideration was dropped from the sample for that regression. In all cases, the relevant number of observations is reported. Thus, for the regression reported in Table 3, 30 of the 112 original countries did not have data for either literacy in 1960 or life expectancy in 1960.

One last preliminary must be noted. Heston and Summers provide four different grades (A to D) that capture their estimate of the quality of the data for different countries. A preliminary least squares regression of growth rates on a trend, investment, and consumption was estimated, and the residuals were checked for evidence of heteroskedasticity related to data quality. The root mean squared residuals were virtually identical for the countries with data of grades B, C, and D and were roughly twice as large (specifically in the ratio of 2.3 to 1.2) as those for the A countries. These results were used to provide weights that were used in all of the subsequent analysis.

With all this as background, it is possible to turn to the table itself.

The growth rates are measured in per cent, so that a 1% average annual growth rate is coded as .01. The literacy and share variables are measured as percent times 100. (Refer to Table 1 for a summary of units and ranges of the variables.) From this information, the magnitudes of the coefficients can be assessed. The estimated coefficient of .00147 for the share of investment in total GDP implies that an increase in the share from 10% to 20% is associated with an increase in the growth rate of .00147 x 10 or 1.47 percent. This number is slightly larger than, but roughly consistent with the magnitude that one would expect from a growth accounting analysis. An increase of 10% in I/Y implies an increases of 3.3% in $\dot{K}/K$ if the capital-output ratio is around one third. If capital's share in total income is around .3, this implies an increase in the growth rate equal to 1 percentage point.

The coefficient of around .00050 on literacy implies that an increase of in literacy equal to 10 percentage points is associated with an increase in the growth rate of one half of a percentage point. Given observed values for literacy ranging form 3% to 98%, the estimated effect of this variable is quite large. This is one case where the use of instrumental variables is quite important. If instead of life expectancy, literacy is used as an instrument, the estimated coefficient on literacy decreases to .00018 and as one would expect, the standard error is smaller (.00008, as opposed to .00014.) When the (log of) per capita consumption of newsprint is used as an instrument, the estimate of .00028 is in between these two estimates, and the standard error is the same as that using life expectancy (.00014).

The other notable feature of this table is that the dummy variable for Africa is relatively small and is not precisely estimated. However, the variable for Latin America is large both in economic terms and in comparison with its standard error.

Thus, one interpretation of these results is that they are consistent with the theory outlined above in the following sense. Capital accumulation has an effect that is slightly larger than, but roughly consistent with the effect one would predict using growth accounting based on market prices. Literacy has a separate, and large, effect on output. This kind of result makes sense if one interprets the relevant applied "research" here as operating at the most primitive level, incremental level. Schmookler (1966) makes a wonderful point about innovation with his discussion of the hundreds of small patentable improvements in horseshoes that took place in the United States right up until the 1920's. This is the kind of "applied research" that one must think of here, the kind done by farmers and tradesmen, not the kind done by scientists in white lab coats. The fact that capital and literacy have separate effects suggests that the cross country variation in the rate of improvement induced by literacy is not too closely correlated with the cross country variation in aggregate capital investment.

Continuing for the moment to take the results from Table 3 at face value, one can go further and ask whether the rate of change of literacy has any additional explanatory power in a regression of this form, as growth accounting would suggest, or whether the level of literacy retains its role when its rate of change is included as well. The answer depends on how seriously one wants to take the problem of measurement error. The most favorable conclusions follow by asserting that while the measured level of literacy might not be comparable across countries, changes in the measured literacy rate between roughly 1960 and 1980 should be comparable across countries. Thus, no instrument is needed for the change in literacy, only for its level. In this case, with life expectancy in 1960 used as an instrument for literacy in 1960, and the change in literacy used as an instrument for itself, the estimated coefficients are .00053 for literacy (with a standard error of .00016) and a coefficient of .00042 for the change in literacy (with a standard error of .00019.) None of the other coefficient estimates change appreciably.

If one has less confidence in the data, one could use newsprint consumption and the change in newsprint consumption as the basic indicator of literacy and use life expectancy and the change in literacy as instruments. The more obvious choice of newsprint consumption as an instrument is probably ill advised because there is a very plausible causal connection between increases in income and increases in newsprint consumption. Thus, errors in newsprint consumption are more likely to be correlated with the errors in the growth rate equation. Of course, one can make a similar case that the change in literacy may be caused by the growth rate of income, so the sense in which the change in literacy is a better instrument is only a relative one.

In any case, using these instruments, the estimated coefficient on (the log of) per capita newsprint consumption in 1960 is .015 (standard error .005) and on the change in this variable between 1960 and 1983 of .011 (standard error .005). To make these coefficients roughly comparable to those for literacy, assume that this variable increases by 10% of its range from a minimum of -4 to a maximum of 3, that is by 0.7. Then the implied increase in growth rates would be around 1% for a change in the initial level and around .77% for an increase in the change between 1960 and 1983, numbers that are roughly twice the comparable estimates given above.

The rain on this sensible parade of results is that the estimated effect of the initial level of income is very large, suspiciously so. When one tries to take account of the likely sources of bias in the estimation of this coefficient, the effects of literacy diminish dramatically. The intuition for this interaction can best be seen from Figure 2. This figure gives a scatter plot of the growth rate of per capita income against the initial level (measured in 1980 dollars). Using the coefficient estimates from Table 3, the solid line also plots the level of growth that is predicted as a function of the initial level of income for a country that has a GDP share of government spending and investment equal to the mean levels in the sample (16% and 14% respectively), but that has a level of literacy that is equal to 0. What the figure shows is that increases in initial level of income are estimated to have a very strong negative effect on growth. Given this estimated effect for the initial level, literacy is the only variable in the equation that varies systematically with the initial level that has a chance to offset the implied negative growth rates for the developed countries.

If one had confidence that the estimated negative effects of the initial level are real, multiple regression analysis would separate out these two effects just as it should. However, there is good reason to believe that the estimated level effect is contaminated by measurement error. Suppose that the basic income accounts data on which Summers and Heston must base all of their estimates have measurement errors that are nontrivial in the initial period. In particular, suppose that for the least developed countries, there was wide variation in the coverage of the income accounts in 1960. Countries that started with narrow coverage that broadened over time as the collection of statistics improved would show an erroneously low level of initial income and an erroneously high rate of growth. These are kinds of problems that Heston and Summers can do nothing about, and the use of RY160 as an instrument for RY260 will do nothing to avoid, since both of the estimates are based on the same raw data. It also seems possible that there are other sources of error arising from the process whereby domestic prices are made internationally comparable.

If one has a separate instrument that can be used for the initial level of income, one can control for measurement error, but the independent variables that are likely to be useful for predicting the initial level of income are the same as the ones that are useful for predicting the initial level of literacy. Thus, one is inevitably forced into the kinds of problems of multicollinearity revealed in Table 4. The first panel removes the insignificant African dummy variable and the quadratic and logarithmic terms in the initial level of income. These three restrictions cause a reduction of the log likelihood (which should be distributed as approximate chi-squared with 3 degrees of freedom) of around 4, a value that is not being significant. The second panel shows what happens when a second instrument, the newsprint consumption variable, is used together with life expectancy and the initial level of income is dropped from the instrument list. The estimated coefficient on literacy goes down to one third of its previous value, and the standard errors for literacy, and the initial level of income increase dramatically, by factors of 20 and 30 respectively. All of the standard errors increase somewhat, partly because of a reduction in the number of countries covered, but these large increases are suggestive of collinearity between that part of the variation in measured literacy and in measured initial income that is that is picked up by the instruments.

One additional piece of information that can be brought to bear here is the estimates of data quality. The literacy variable was removed, and four separate coefficients were estimated for the initial level of income, one for each level of data quality, using in each case the initial level of income as its own instrument. Consistent with the idea the that the negative bias in the coefficient will be larger the lower is the quality of the data, the (negative) estimated coefficient on the initial level of per capita income is monotonically decreasing (that is increasing in absolute value) with decreases in the quality of the data, with a ratio between the coefficient for the class A countries and the class D countries that is on the order of 5. However, these coefficients are not very precisely estimated; the marginal significance level of the hypothesis that they are all the same is around .9. Moreover, since data quality is closely related to the initial level of income, this variation cannot be distinguished from the hypothesis that the effect of the initial level of income has a positive curvature, i.e. a positive a quadratic term such as that found by Barro.

Tables 5 and 6 illustrate a related interaction between the variables that is problematic. Table 5 gives information about the variation between the three measured shares of GDP. The first panel gives results for the shares measured in current value terms. The second gives results for shares measured using 1980 price weights. Two features are noteworthy. First, there is much more unexplained variation in 1980 price data than in the current price data. It could be that the true standard deviation in net exports (implied here by the variation in the residual from this equation) is on the order of 10% as in Panel 2 rather than 4% as in Panel 1, but it is more likely that the difficulties inherent in using fixed year prices lead to substantial measurement problems. This is suggested further by the fact that the coefficients in the first panel are more plausible. Together, these offer some support for the prior assertion that current value quantities are likely to be more appropriate for the purposes here.

The second noteworthy feature is that even in the first panel, the share of consumption does not respond one for with changes in the share of government spending when the share of investment is held constant. To interpret this finding, it is useful to rewrite the equation

$$\text{CONS} = \text{C} + \alpha\ \text{CGOV} + \beta\ \text{CINV} + \epsilon \tag{11}$$

as

$$\begin{aligned}\text{NET EXPORTS} &= 100 - \text{CONS} - \text{CGOV} - \text{CINV} \\ &= 100 - \text{C} - (1-\alpha)\ \text{CGOV} - (1-\beta)\ \text{CINV} - \epsilon \end{aligned} \tag{12}$$

The estimate of C is very close to 100 and the coefficient β on CINV is close enough to 1 to ignore the difference. But α is far from 1, implying a negative relation between net exports and government spending as shares of GDP. It is not clear what the source of this relation is. For the poorest countries, one candidate explanation is direct foreign aid and grants that are at least partially counted as government spending. Consistent with this view is the finding that the size of the absolute value of the implied residuals from these equations is monotonically related to the estimate of the quality of the data, with the D countries having the largest residuals. Moreover, the residuals from the more plausible equation in Panel 1 are on average negative for the countries with data grades A, B and C with a value of around −0.3 (implying positive net exports of 0.3% of GDP) but are positive with a value of 1.2 for the D countries (implying net exports of −1.2% of GDP for these countries on average.) The finding that the size of the absolute value of the residuals increases as data quality decreases is consistent with pure measurement error in the data, but the finding that the sign of the residuals varies with the data quality is suggestive of a role for transfers.

The primary theoretical rationale for a negative effect of government spending on growth is one that operates through the incentive effects of distortionary taxation. (There are subtleties here about whether taxes should still have a role if accumulation is measured directly as it is here through investment. For taxes to have separate role, it must be the case that they limit accumulation of inputs that are not adequately measured in investment.) To the extent that increases in current government expenditures do not lead to reductions in current consumption (or to expected future reductions in consumption), they should not have a negative effect on growth. Thus, one can think of measured government spending as being that part of spending financed by distortionary taxes plus an error term that is not correlated with current consumption. Thus, consumption can be used as an instrument for government spending, and when it is, one would expect to find an increase in the absolute magnitude of the coefficient on CGOV; that is, it should become more negative.

Table 6 shows what in fact happens when CONS is used an an instrument for CGOV in regressions that include the literacy variable LT60. The table repeats the two regressions from Table 4, substituting CONS for GOV in the instrument list. One interpretation of these results is that CONS is just a bad instrument for CGOV. It makes little difference in the first regression and everything deteriorates dramatically when it is used in the second. An alternative interpretation is that there are two changes that bring out problems with collinearity: removing that part of CGOV that is not correlated with changes in CONS and INV, and using an instrument to remove the bias in the estimates of the coefficient on RY260. When the part of CGOV that is not correlated with CONS and INV is taken away in moving from Table 4 to Table 6, the standard error of CGOV increases by a factor of 3 in Panel 1, and by a factor of 5 in Panel 2. In the second panel, the sign of the coefficient also switches. A comparison of Panels 1 and 2 in both Tables 4 and 6 shows the effect of using an instrument for RY260 that avoids the measurement error bias. Doing so increases the standard error on LT60 and RY260 by an order of magnitude as was noted above. The nebulous results reported in Panel 2 of Table 6 suffer from both of these effects.

If collinearity is indeed part of the problem, excluding one or the other of CGOV, LT60, or RY260 should reduce the standard errors of the estimates considerably. Table 7 repeats the regression from Panel 2 of Table 6, excluding CGOV in Panel 1, excluding LT60 in Panel 2, and excluding RY260 in panel 3. In the first panel, the effect of excluding CGOV is not impressive. The coefficient on LT60 retains the implausibly high value it held in Panel 2 of Table 6, more than 5 times its previously estimated value. It implies that an increase in literacy from the smallest value of 1 to the largest value of 99 would cause a difference in growth rates equal to 14 percentage points. Its standard error also remains very high.

When literacy is removed in the second panel, the standard errors on both CGOV and CINV are smaller than those in panel 2 of Table 6, falling in the first case by a factor of 8, in the second case by a factor of 2. The coefficient on investment takes on a value in the upper end of the range of values noted so far, one that is about twice what one would expect based on the simple growth accounting calculation given above if this coefficient is interpreted as the causal effect from exogenous changes in investment. The coefficient on the share of government is also quite large. Over the observed range of values of CGOV from 5 to 35, this coefficient implies a change in growth rates of 9 percentage points if it is given a causal interpretation. In this regression, it also is possible to retain the newsprint consumption variable as an additional instrument. This has the effect of reducing the number of observations back down to 66 as in panel 1 of the table. This has little effect on the qualitative conclusions described here. In particular, the standard errors are smaller than in panel 1, even with the smaller number of observations.

Panel 3 shows that excluding the initial level of income has much the same effect as excluding literacy. Compared to the regression in panel 2 of Table 6 in which all the variables are included, the standard errors are lower and the estimated coefficients on investment and the share of government are larger.

The main finding from these regressions is that although the standard errors are reduced when a variable is omitted, neither the initial level of income nor the initial level of literacy has an estimated coefficient in any of these regressions.

Table 8 shows that the much larger estimate of the effect of the share of government described in the last two regressions is attributable almost entirely to the use of CONS as an instrument, and not to the exclusion of literacy or of the initial level of income. This table repeats the last two regressions using GOV as the instrument for CGOV instead of CONS. Just as one would expect from the use of an instrumental variables estimate when measurement error is present, the standard error in Panel 2 is larger, but the coefficient is also larger, in this case, very much so.

Tables 9 and 10 conclude the diagnostic checks by reporting the first stage regressions for literacy and the initial level of income. The key observation here is that the R^2 statistics are each case agreeably high. The problem here is not bad instruments. These give further evidence that the ambiguous results reported in the Tables 4 and 6 are not just due to the fact that the instruments are bad. Taken together with the evidence from the regressions that exclude the different variables, these offer strong evidence that the fundamental problem in those tables is multicollinearity, especially between the initial level of income and the initial level of literacy, that is uncovered when a correction for measurement error in the initial level of income is used.

4. Conclusion

The empirical results are summarized in the introduction, and there is no reason to repeat this summary here. As has already been noted, the results here are only the beginning of the consideration of these data in the light of the kind of model outlined here. The support for a direct role for literacy in increasing growth rates is tenuous at best, but the model suggests that this might be the case if investment is one of the other variables that is taken as given. The next steps are to investigate the effect of the initial level of literacy on investment, and to explore the role of measures of the advanced human capital like scientific and engineering talent. Preliminary explorations of these issues appear to be supportive of the model. The initial level of literacy does seem to be significantly related to investment even when other variables are held constant. Measures of scientific talent seem to be positively related to both growth and investment in the small sample of developed countries where it is present in any appreciable quantity.

At a methodological level, the major conclusion here is a sobering one, but it need not be a discouraging one. As one should have suspected given the underlying sources, the cross country data seem to be subject to measurement error, but this does not mean that there is nothing that can be learned from them. On the contrary, there appears to be much that can be learned. Because there are so many different indicators of the same underlying variables, there is real hope that the measurement errors can be overcome. One can only hope that someone will someday put as much effort into organizing the collateral data from the UNESCO and the World Bank as Summers and Heston have devoted to organizing the national income accounts. Together, these sources should prove quite revealing to economists who are willing to proceed with a measure of caution.

References

Arrow, K. J. 1962a. "The economic implications of learning by doing." **Review of Economic Studies**, 29 (June), p. 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. "Threshold Externalities in Economic Development." University of Pennsylvania Working paper, October 1988.

Barro, R. J. 1989. "A Cross Country Study of Growth, Savings, and Government." Harvard University, December 1988.

Becker, G. and K. Murphy. 1988. "Economic Growth, Human Capital, and Population Growth." University of Chicago, June 1988.

Chamberlain, G. and Z. Griliches. "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, p. 97–124.

Grossman, G and E. Helpman. "Comparative advantage and long run growth." Princeton University, June 1988.

Hicks, N. 1979. "Growth vs. basic needs: Is there a tradeoff?". World Development 7 (November/December) p. 985–94.

——————. 1980. "Is there a tradeoff between growth and basic needs." **Finance and Development** (June), p. 17–20.

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

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

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." NBER Macroeconomics Annual, S. Fischer ed. Cambridge: MIT PRess.

————. 1988. "Endogenous Technological Change." University of Chicago (May).

————. 1989. "Capital accumulation in the theory of long run growth." In

Schmookler, J.. 1962. Invention and Economic Growth. Cambridge: Harvard University Press.

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

Solow, R. 1956. "A contribution to the theory of economic growth." Quarterly Journal of Economics 70 (February), p. 65–94.

Summers, R. and A. 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) p. 1–25.

Wilson, R. 1975. "Informational Economies of Scale." Bell Journal of Economics 6 (Spring) p. 184–195.

Table 1: Variable Definitions

C	A constant term used in all of the regressions.
RY260	Real per capita income in 1960 in current 1980 prices, using current price weights. From Summers and Heston, RGDP2. Range, $250 to $7400.
RY2G	The average anual rate of growth in percent of RY2 over the years 1960 to 1985. Range, -.04 to .07.
RY160	Real per capita income in 1960 measured in 1980 prices, using 1980 price weights. From Summers and Heston, RGDP1. Same range as RY260.
CGOV, GOV	Share of GDP government spending on items other than investment goods, in percent times 100, avearaged over the years 1960 to 1985. CGOV measured using current price weights, GOV using 1980 price weights. Range: 5 to 35.
CINV, INV	Share of GDP devoted to investment, averaged over 1960 to 1985. CINV measured using current price weights, INV using 1980 price weights. Range, 4 to 37.
CONS	Share of GDP devoted to consumption, averaged over 1960 to 1985. Measured using 1980 price weights. Range 25 to 104.
LT60	Percent of the population times 100 that is literate in survey year close to 1960. Range, 1 to 98.
NP60L	The logarithm of the per capita comsumption of newsprint in 1960. Range, -4 to 4.
EX60	Life expectancy in years in 1960 and 1986. Range 40 to 70.

Codes: As indicated above, the prefix letter "C" denotes a current price version of a variable. A suffix letter "L" denotes the log of a variable. The suffix letter "Q", used only for the RY1 and RY2 variables denotes a quadratic term.

Table 2: Countries: Names, Numbers and Data Grades from Summers and Heston

1	Algeria	C
2	Angola	D
3	Benin	D
4	Botswana	B
6	Burundi	D
7	Cameroon	C
8	Central Afr. Rep.	D
9	Chad	D
10	Congo, Peop. Rep.	D
11	Egypt	D
12	Ethiopia	C
13	Gabon	C
14	Gambia, The	D
15	Ghana	D
16	Guinea	D
17	Ivory Coast	C
18	Kenya	B
19	Lesotho	D
20	Liberia	D
21	Madagascar	C
22	Malawi	C
23	Mali	C
24	Mauritania	D
25	Mauritius	D
26	Morocco	C
27	Mozambique	D
28	Niger	D
29	Nigeria	C
30	Rwanda	D
31	Senegal	C
32	Sierra Leone	D
33	Somalia	D
34	S. Africa	B
35	Sudan	D
36	Swaziland	D
37	Tanzania	C
38	Togo	D
39	Tunisia	C
40	Uganda	D
41	Zaire	D
42	Zambia	B
43	Zimbabwe	C
46	Bangladesh	C
47	Burma	C
48	Hong Kong	A
49	India	B
50	Iran	C
51	Iraq	C
52	Israel	A
53	Japan	A

Table 2 (cont)

54	Jordan	C
55	Korea, Rep. of	B
57	Malaysia	B
58	Nepal	D
60	Pakistan	B
61	Philippines	A
63	Singapore	C
64	Sri Lanka	B
65	Syrian Arab Rep.	C
66	Taiwan	B
67	Thailand	C
70	Austria	A
71	Belgium	A
72	Cyprus	B
73	Denmark	A
74	Finland	A
75	France	A
76	Germany, Fed Rep	A
77	Greece	A
78	Iceland	B
79	Ireland	A
80	Italy	A
81	Luxembourg	A
82	Malta	B
83	Netherlands	A
84	Norway	A
85	Portugal	A
86	Spain	A
87	Sweden	A
88	Switzerland	B
89	Turkey	B
90	United Kingdom	A
91	Barbados	C
92	Canada	A
93	Costa Rica	B
94	Dominican Rep.	C
95	El Salvador	B
96	Guatemala	B
97	Haiti	C
98	Honduras	C
99	Jamaica	C
100	Mexico	B
101	Nicaragua	C
102	Panama	B
103	Trinidad & Tobago	C
104	United States	A
105	Argentina	B
106	Bolivia	B
107	Brazil	B
108	Chile	C
109	Colombia	B

Table 2 (cont.)

110	Ecuador	B
111	Guyana	C
112	Paraguay	C
113	Peru	C
114	Surinam	C
115	Uruguay	B
116	Venezuela	B
117	Australia	A
118	Fiji	C
120	New Zealand	A
121	Papua New Guinea	D

Table 3

TSLS // Dependent Variable is RY2G
Number of observations: 82
Instrument list: C RY160 RY160Q RY160L GOV INV LADUM AFDUM EX60
Weighting series: WT

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.0009989	0.0574736	-0.0173801	0.986
RY260	-1.828E-05	1.305E-05	-1.4003002	0.166
RY260Q	1.093E-09	1.522E-09	0.7180084	0.475
RY260L	0.0025460	0.0098475	0.2585374	0.797
CGOV	-0.0008957	0.0003676	-2.4367744	0.018
CINV	0.0014712	0.0003878	3.7935532	0.000
LADUM	-0.0144710	0.0052102	-2.7774643	0.007
AFDUM	-0.0066555	0.0056562	-1.1766731	0.244
LT60	0.0004962	0.0001461	3.3971464	0.001

Weighted Statistics

R-squared	0.470606	Mean of dependent var	0.014849
Adjusted R-squared	0.412590	S.D. of dependent var	0.019092
S.E. of regression	0.014632	Sum of squared resid	0.015630
Durbin-Watson stat	2.472103	F-statistic	8.111685
Log likelihood	234.8237		

Unweighted Statistics

R-squared	0.537848	Mean of dependent var	0.016188
Adjusted R-squared	0.487201	S.D. of dependent var	0.019943
S.E. of regression	0.014281	Sum of squared resid	0.014888
Durbin-Watson stat	2.575129		

Table 4: TSLS, Dependent variable RY2G

Panel 1: RY160 and EX60 used as instruments for RY260 and LT60

Number of observations: 82
Instrument list: C RY160 GOV INV LADUM EX60

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0076044	0.0076326	0.9963140	0.323
RY260	-1.120E-05	2.763E-06	-4.0531803	0.000
CGOV	-0.0010047	0.0003690	-2.7223970	0.008
CINV	0.0014208	0.0003715	3.8243884	0.000
LADUM	-0.0140912	0.0050946	-2.7658897	0.007
LT60	0.0005557	0.0001336	4.1591051	0.000

Panel 2: EX60 and NP60L used as instruments for RY260 and LT60

Number of observations: 66
Instrument list: C EX60 GOV INV LADUM NP60L

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0174817	0.0435753	0.4011841	0.690
RY260	-4.974E-05	8.710E-05	-0.5710753	0.570
CGOV	-0.0015920	0.0022310	-0.7135931	0.478
CINV	0.0018898	0.0010530	1.7946638	0.078
LADUM	-0.0252761	0.0348497	-0.7252910	0.471
LT60	0.0016266	0.0026578	0.6120040	0.543

Table 5: Least Square Regression, Dependent Variable CCONS, CONS

Panel 1: Shares measured in current prices

Dependent Variable CCONS
Number of observations: 112

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	99.679519	1.5999203	62.302803	0.000
CGOV	-0.7645585	0.0736063	-10.387133	0.000
CINV	-1.0790575	0.0619481	-17.418728	0.000

R-squared	0.779664	Mean of dependent var	71.43429
Adjusted R-squared	0.775622	S.D. of dependent var	8.800402
S.E. of regression	4.168627	Sum of squared resid	1894.142
Durbin-Watson stat	2.372530	F-statistic	192.8500
Log likelihood	-317.2904		

Panel 2: Shares measured in 1980 prices

Dependent Variable CONS
Number of observations: 112

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	85.154500	3.7529394	22.690081	0.000
GOV	-0.1666855	0.1515545	-1.0998384	0.274
INV	-0.8431423	0.1225247	-6.8814066	0.000

R-squared	0.304955	Mean of dependent var	66.68591
Adjusted R-squared	0.292202	S.D. of dependent var	12.08443
S.E. of regression	10.16672	Sum of squared resid	11266.49
Durbin-Watson stat	1.806548	F-statistic	23.91217
Log likelihood	-417.1421		

Table 6: TSLS, Dependent variable RY2G

Panel 1: Repeat Panel 1 of Table 4 using CONS as the instrument for CGOV

Number of observations: 82
Instrument list: C RY160 CONS INV LADUM EX60

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0132727	0.0251584	0.5275643	0.600
RY260	-1.114E-05	2.736E-06	-4.0731040	0.000
CGOV	-0.0013426	0.0015109	-0.8886074	0.377
CINV	0.0015252	0.0006723	2.2686386	0.026
LADUM	-0.0139290	0.0052081	-2.6744763	0.009
LT60	0.0005120	0.0002564	1.9971462	0.050

Panel 2: Repeat Panel 2 of Table 4 using CONS as the instrument for CGOV

Number of observations: 66
Instrument list: C EX60 CONS INV LADUM NP60L

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.0586466	0.1097875	-0.5341825	0.595
RY260	-2.720E-05	2.587E-05	-1.0514961	0.297
CGOV	0.0026833	0.0062242	0.4311065	0.668
CINV	0.0013463	0.0011798	1.1411263	0.258
LADUM	-0.0156504	0.0120429	-1.2995567	0.199
LT60	0.0012787	0.0013629	0.9382146	0.352

Table 7: TSLS Estimates, Dependent Variable RY2G

Panel 1: CGOV excluded. Otherwise, same as Table 6, Panel 2.

Number of observations: 66
Instrument list: C EX60 INV LADUM NP60L

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.0108665	0.0113918	-0.9538888	0.344
RY260	-4.135E-05	6.580E-05	-0.6284118	0.532
CINV	0.0016874	0.0008931	1.8894373	0.064
LADUM	-0.0216917	0.0263203	-0.8241442	0.413
LT60	0.0014970	0.0021301	0.7028057	0.485

Panel 2: LT60 excluded. Otherwise, same as Table 6, Panel 2.

Number of observations: 97
Instrument list: C EX60 CONS INV LADUM

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0444133	0.0135638	3.2743934	0.002
RY260	-2.762E-06	2.905E-06	-0.9507838	0.345
CGOV	-0.0030514	0.0008245	-3.7010088	0.000
CINV	0.0019617	0.0004860	4.0362520	0.000
LADUM	-0.0094833	0.0047978	-1.9765825	0.052

Panel 3: RY260 excluded. Otherwise, same as Table 6, Panel 2.

Number of observations: 66
Instrument list: C EX60 CONS INV LADUM NP60L

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0464978	0.0272532	1.7061406	0.093
CGOV	-0.0032603	0.0015668	-2.0808966	0.042
CINV	0.0020909	0.0005674	3.6848548	0.000
LADUM	-0.0082190	0.0058632	-1.4017972	0.166
LT60	-0.0001096	0.0002032	-0.5394303	0.592

Table 8: TSLS Estimates, Dependent Variable RY2G

Panel 1: GOV instead of CONS as instrument. Otherwise, same as panel 2 Table 7, with LT60 excluded.

Number of observations: 97
Instrument list: C EX60 GOV INV LADUM

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	0.0067213	0.0075949	0.8849723	0.379
RY260	3.299E-06	2.286E-06	1.4427895	0.153
CGOV	-0.0007485	0.0003663	-2.0430312	0.044
CINV	0.0012724	0.0003903	3.2600762	0.002
LADUM	-0.0037635	0.0041614	-0.9043853	0.369

Panel 2: GOV instead of CONS as instrument. Otherwise, same as panel 3 Table 7, with RY260 excluded.

Number of observations: 66
Instrument list: C GOV INV LADUM EX60 NP60L

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.0022107	0.0099579	-0.2220092	0.825
CGOV	-0.0005381	0.0004685	-1.1483938	0.255
CINV	0.0018038	0.0003896	4.6304739	0.000
LADUM	-0.0068938	0.0049920	-1.3809565	0.172
LT60	0.0001174	0.0001055	1.1125596	0.270

Table 9: First Stage Regression, Dependent variable LT60

Number of observations: 66

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-49.329702	18.404700	-2.6802774	0.009
CONS	0.1316173	0.1460933	0.9009133	0.371
INV	-0.4035904	0.3256701	-1.2392614	0.220
LADUM	8.8027638	3.9257055	2.2423393	0.029
EX60	1.8194340	0.3136788	5.8003092	0.000
NP60L	6.2883467	1.8300380	3.4361837	0.001

Weighted Statistics

R-squared	0.797965	Mean of dependent var	45.36940
Adjusted R-squared	0.781129	S.D. of dependent var	27.27787
S.E. of regression	12.76159	Sum of squared resid	9771.496
Durbin-Watson stat	2.256653	F-statistic	47.39564
Log likelihood	-258.5698		

Unweighted Statistics

R-squared	0.846956	Mean of dependent var	48.40000
Adjusted R-squared	0.834202	S.D. of dependent var	30.33477
S.E. of regression	12.35180	Sum of squared resid	9154.015
Durbin-Watson stat	2.319464		

Table 10: First Stage Regression, RY260 dependent variable.

Number of observations: 79

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	642.35561	1136.9977	0.5649577	0.574
CONS	-25.348800	9.0200210	-2.8102817	0.007
INV	-40.416349	19.688656	-2.0527734	0.044
LADUM	-513.27234	227.42066	-2.2569292	0.027
EX60	68.851163	19.176904	3.5903170	0.001
NP60L	362.85153	110.50065	3.2837051	0.002

Weighted Statistics

R-squared	0.602174	Mean of dependent var	1777.547
Adjusted R-squared	0.574925	S.D. of dependent var	1302.729
S.E. of regression	849.3503	Sum of squared resid	52661905
Durbin-Watson stat	2.419251	F-statistic	22.09943
Log likelihood	-641.7894		

Unweighted Statistics

R-squared	0.749050	Mean of dependent var	2067.380
Adjusted R-squared	0.731862	S.D. of dependent var	1782.595
S.E. of regression	923.0640	Sum of squared resid	62199443
Durbin-Watson stat	1.789526		

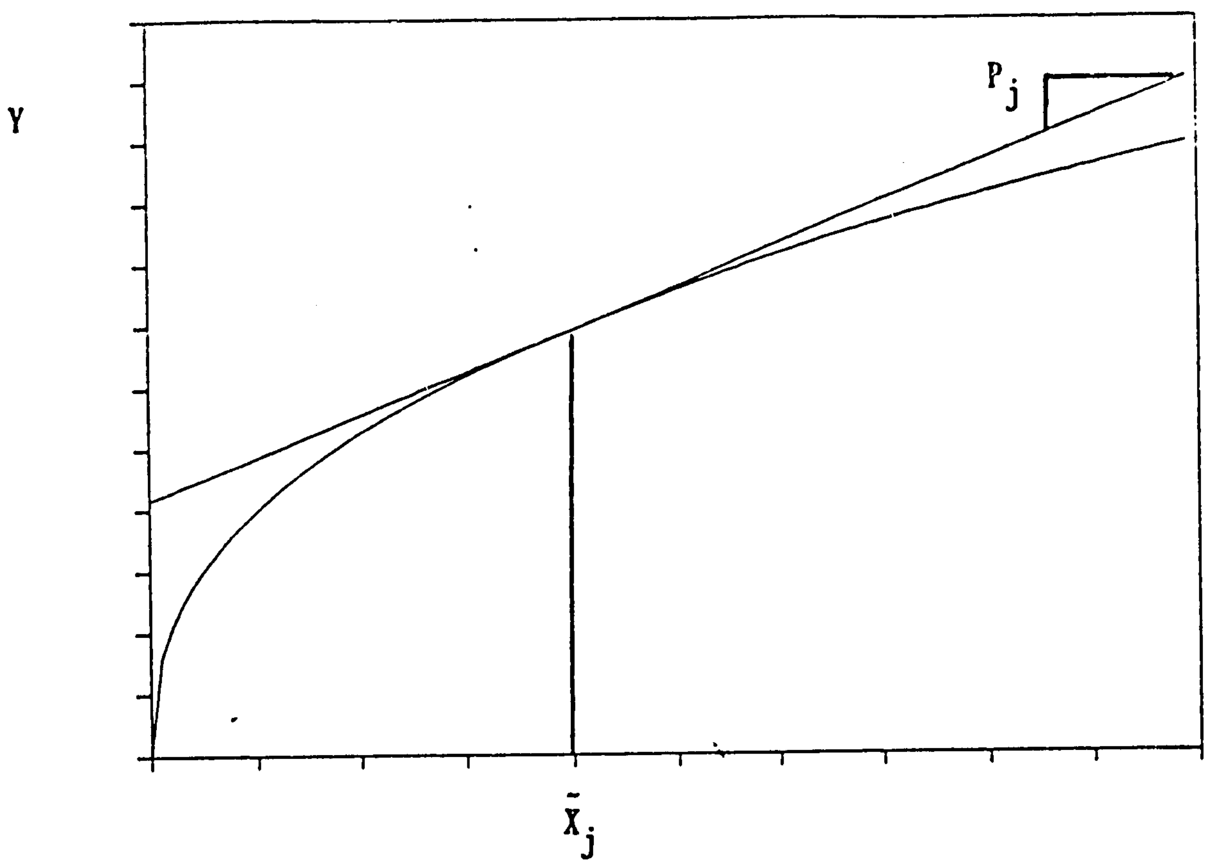
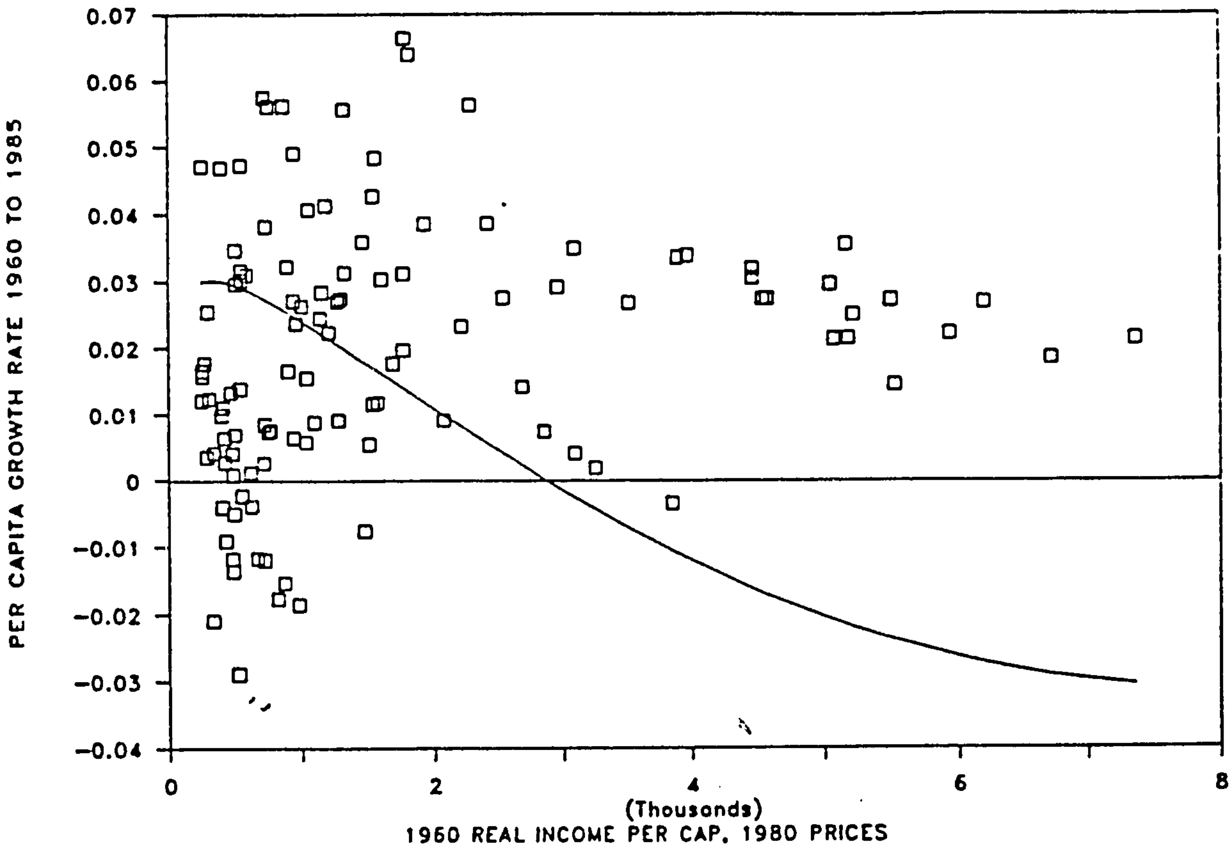


Figure 1



——— = Predicted Growth from Regression in Table 3 based on Mean CGOV and CINV in the Sample and LT60 = 0.

Figure 2