ANNALS OF ECONOMICS AND FINANCE **14-2(A)**, 305–342 (2013)

Health and Economic Growth

Robert J. Barro

Harvard University

1. INTRODUCTION

Since the mid 1980s, research on economic growth has experienced a boom, beginning with the work of Romer (1986). The new "endogenous growth" theories have focused on productivity advances that derive from technological progress and increased human capital in the form of education. Barro and Sala-i-Martin (1995) explore these theories and also discuss extensions to allow for open economies, diffusion of technology, migration of persons, fertility choice, and variable labor supply. The government can be important in the models in terms of its policies on maintenance of property rights, encouragement of free markets, taxation, education, and public infrastructure.

One area that has received little attention in the recent literature on growth theory is the two-way interplay between health and economic growth. Two preliminary efforts in this direction are Ehrlich and Lui (1991) and Meltzer (1995). Also, the empirical work of Barro (1996) and others suggests that health status, as measured by life expectancy or analogous aggregate indicators, is an important contributor to subsequent growth. In fact, initial health seems to be a better predictor than initial education of subsequent economic growth.

The main purpose of this study is to apply the spirit and apparatus of the recent advances in growth theory to the interaction between health and growth. The analysis is conceptual and is intended to form the basis for further theorizing and for empirical analyses of the joint determination of health and growth.

The discussion begins with a survey of existing theories and empirical evidence on the determinants of economic growth. Then the paper develops models of the interplay between health and growth.

1529-7373/2013

2. A SUMMARY OF THEORY AND EVIDENCE ON ECONOMIC GROWTH

2.1. Old and New Theories of Economic Growth

In the 1960s, growth theory consisted mainly of the neoclassical model, as developed by Ramsey (1928), Solow (1956), Swan (1956), Cass (1965), and Koopmans (1965). One feature of this model, which has been exploited seriously as an empirical hypothesis only in recent years, is the convergence property. The lower the starting level of real per capita gross domestic product (GDP) the higher is the predicted growth rate.

If all economies were intrinsically the same, except for their starting capital intensities, then convergence would apply in an absolute sense; that is, poor places would tend to grow faster per capita than rich ones. However, if economies differ in various respects — including propensities to save and have children, willingness to work, access to technology, and government policies — then the convergence force applies only in a conditional sense. The growth rate tends to be high if the starting per capita GDP is low in relation to its long-run or steady-state position; that is, if an economy begins far below its own target position. For example, a poor country that also has a low long-term position — possibly because its public policies are harmful or its saving rate is low — would not tend to grow rapidly.

The convergence property derives in the neoclassical model from the diminishing returns to capital. Economies that have less capital per worker (relative to their long-run capital per worker) tend to have higher rates of return and higher growth rates. The convergence is conditional because the steady-state levels of capital and output per worker depend in the neoclassical model on the propensity to save, the growth rate of population, and the position of the production function — characteristics that may vary across economies. Recent extensions of the model suggest the inclusion of additional sources of cross-country variation, especially government policies with respect to levels of consumption spending, protection of property rights, and distortions of domestic and international markets.

The concept of capital in the neoclassical model can be usefully broadened from physical goods to include human capital in the forms of education, experience, and health. (See Lucas (1988), Rebelo (1991), Caballe and Santos (1993), Mulligan and Sala-i-Martin (1993), and Barro and Sala-i-Martin (1995a, Ch. 5).) The economy tends toward a steady-state ratio of human to physical capital, but the ratio may depart from its long-run value in an initial state. The extent of this departure generally affects the rate at which per capita output approaches its steady-state value. For example, a country that starts with a high ratio of human to physical capital (perhaps because of a war that destroyed mainly physical capital) tends to grow rapidly because physical capital is more amenable than human capital to rapid expansion. A supporting force is that the adaptation of foreign technologies is facilitated by a large endowment of human capital (see Nelson and Phelps (1966) and Benhabib and Spiegel (1994)). This element implies an interaction effect whereby a country's growth rate is more sensitive to its starting level of per capita output the greater is its initial stock of human capital.

Another prediction of the neoclassical model — even when extended to include human capital — is that, in the absence of continuing improvements in technology, per capita growth must eventually cease. This prediction, which resembles those of Malthus (1798) and Ricardo (1817), comes from the assumption of diminishing returns to a broad concept of capital. The long-run data for many countries indicate, however, that positive rates of per capita growth can persist over a century or more and that these growth rates have no clear tendency to decline.

Growth theorists of the 1950s and 1960s recognized this modeling deficiency and usually patched it up by assuming that technological progress occurred in an unexplained (exogenous) manner. This device can reconcile the theory with a positive, possibly constant per capita growth rate in the long run, while retaining the prediction of conditional convergence. The obvious shortcoming, however, is that the long-run per capita growth rate is determined entirely by an element — the rate of technological progress — that comes from outside of the model. (The long-run growth rate of the level of output depends also on the growth rate of population, another element that is exogenous in the standard theory.) Thus, we end up with a model of growth that explains everything but long-run growth, an obviously unsatisfactory situation.

Recent work on endogenous growth theory has sought to supply the missing explanation of long-run growth. In the main, this approach provides a theory of technical progress, one of the central missing elements of the neoclassical model. The new models operate by including incentives for the private sector to carry out the research that leads to discoveries of new products or methods of production. Typically, the private reward for invention features elements of monopoly profits over some interval. Patent protection and intellectual property rights affect these private incentives, but the government can also influence research through public subsidies or direct participation. This general framework for technological advance applies, in particular, to discoveries of medicines or medical procedures.

The initial wave of the new research — Ramer (1986), Lucas (1988), Rebelo (1991) — built on the work of Arrow (1962), Sheshinski (1967), and Uzawa (1965) and did not really introduce a theory of technological change. In these models, growth may go on indefinitely because the returns to investment in a broad class of capital goods, which includes human capital, do not necessarily diminish as economies develop. (This idea goes back to Knight (1944)) Spillovers of knowledge across producers and external benefits from human capital are parts of this process, but only because they help to avoid the tendency for diminishing returns to capital.

The incorporation of R&D theories and imperfect competition into the growth framework began with Romer (1987, 1990) and includes significant contributions by Aghion and Howitt (1992) and Grossman and Helpman (1991, Chapters 3 and 4). Barro and Sala-i-Martin (1995, Chs. 6, 7) provide expositions and extensions of these models. In these settings, technological advance results from purposive R&D activity, and this activity is rewarded, along the lines of Schumpeter (1934), by some form of ex-post monopoly power. If there is no tendency to run out of ideas, then growth rates can remain positive in the long run. The rate of growth and the underlying amount of inventive activity tend, however, not to be Pareto optimal because of distortions related to the creation of the new goods and methods of production. In these frameworks, the long-term growth rate depends on governmental actions, such as taxation, maintenance of law and order, provision of infrastructure services, protection of intellectual property rights, and regulations of international trade, financial markets, and other aspects of the economy. The government therefore has great potential for good or ill through its influence on the long-term rate of growth.

One shortcoming of the early versions of endogenous growth theories is that they no longer predicted conditional convergence. Since this behavior is a strong empirical regularity in the data for countries and regions, it was important to extend the new theories to restore the convergence property. One such extension involves the diffusion of technology. Whereas the analysis of discovery relates to the rate of technological progress in leading-edge economies, the study of diffusion pertains to the manner in which follower economies share by imitation in these advances. Since imitation tends to be cheaper than innovation, the diffusion models predict a form of conditional convergence that resembles the predictions of the neoclassical growth model. Therefore, this framework combines the long-run growth of the endogenous growth theories (from the discovery of ideas in the leading-edge economies) with the convergence behavior of the neoclassical growth model (from the gradual imitation by followers).

Endogenous growth theories that include the discovery of new ideas and methods of production are important for providing possible explanations for long-term growth. Yet the recent cross-country empirical work on growth has received more inspiration from the older, neoclassical model, as extended to include government policies, investments in human capital, fertility choice, and the diffusion of technology. Theories of basic technological change seem most important for understanding why the world as a whole can continue to grow indefinitely in per capita terms. But these theories have less to do with the determination of relative rates of growth across countries, the key element studied in the cross-country empirical work that is discussed next.

2.2. Empirical Framework for the Analysis of Growth Across Countries

A standard framework for the determination of growth follows the extended version of the neoclassical model as already described. In equation form, the model can be represented as

$$Dy = f(y, y^*), \tag{1}$$

where Dy is the growth rate of per capita output, y is the current level of per capita output, and y^* is the long-run or steady-state level of per capita output.[1] The growth rate, Dy, is diminishing in y for given y^* and rising in y^* for given y. The target value y^* depends on an array of choice and environmental variables. The private sector's choices involve saving, labor supply, investments in schooling and health, and fertility rates, each of which depends on preferences and costs. The government's choices involve spending in various categories, notably infrastructure, schooling, and public health; tax rates; the extent of distortions of markets and business decisions; maintenance of the rule of law and property rights; and the degree of political freedom. Also relevant for an open economy is the terms of trade, typically given to a small country by external conditions.

For a given initial level of per capita output, y, an increase in the steady-state level, y^*, raises the per capita growth rate over a transition interval. For example, if the government improves the climate for business activity — say by reducing the burdens from regulation, corruption, and taxation, or by enhancing property rights — the growth rate increases for awhile. Similar effects arise if people decide to have fewer children or (at least in a closed economy) to save a larger fraction of their incomes.

In these cases, the increase in the target, y^*, translates into a transitional increase in the economy's growth rate. As output, y, rises, the workings of diminishing returns eventually restore the growth rate, Dy, to a value determined by the rate of technological progress. Since the transitions tend to be lengthy, the growth effects from shifts in government policy or private behavior persist for a long time.

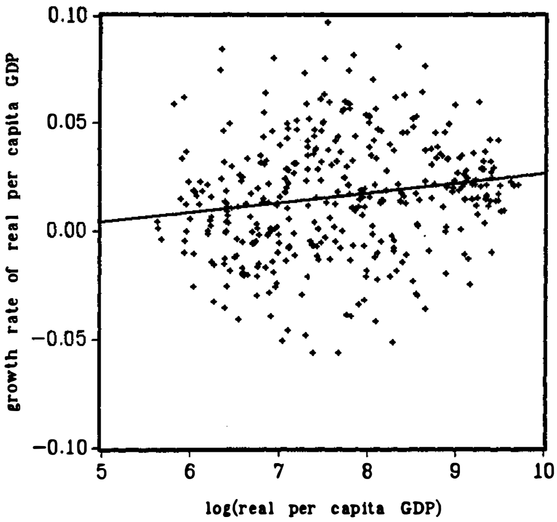
For given values of the choice and environmental variables — and, hence, y^* — a higher starting level of per capita output, y, implies a lower per capita growth rate. This effect corresponds to conditional convergence. Note, however, that poor countries would not grow rapidly on average if

[1] With exogenous, labor-augmenting technological progress, the level of output per worker grows in the long run, but the level of output per effective worker approaches a constant, y^*. Hence, y^* should be interpreted in this generalized sense.

they tend also to have low steady-state positions, y^*. In fact, a low level of y^* explains why a country would typically have a low observed value of y in some arbitrarily chosen initial period.

The last result shows that the framework can be reconciled with the now familiar lack of correlation between the growth rate and initial level of real per capita GDP across a large number of countries over the period 1960 to 1990. Figure 1 shows that this relationship is virtually nil.[2] (The slope actually has the wrong sign — slightly positive — but is not statistically significant.) The interpretation from the standpoint of the neoclassical model is that the initially poor countries, which show up closer to the origin along the horizontal axis, are not systematically far below their steady-state positions and therefore do not tend to grow relatively fast. The isolation of the convergence force requires a conditioning on the determinants of the steady state, as in the cross-country empirical analysis discussed in the next section.

FIG. 1. Simple Correlation between Growth and Level of GDP



2.3. Empirical Findings on Growth across Countries

Table 1 shows results from regressions that use the general framework of equation (1) from the previous section. The regressions apply to a panel

[2]The data on real per capita GDP are the internationally comparable values generated by Summers and Heston (1993). The vertical axis in Figure 1 contains observations on per capita growth rates for 1965-75, 1975-85, and 1985-90, the three periods used in the detailed empirical analysis described below. The horizontal axis shows the corresponding values of the logarithm of per capita GDP in 1965, 1975, and 1985.

of roughly 100 countries observed from 1960 to 1990.[3] The dependent variables are the growth rates of real per capita GDP over three periods: 1965-75, 1975-85, and 1985-90.[4] (The first period begins in 1965, rather than 1960, so that the 1960 value of real per capita GDP can be used as an instrument; see below.) Henceforth, the term GDP will be used as a shorthand to refer to real per capita GDP.

Some previous analysis, such as Barro (1991), used a cross-sectional framework; that is, the growth rate and the explanatory variables were observed only once per country. The main reason to extend to a panel setup is to expand the sample information. Although the main evidence turns out to come from the cross-sectional (between-country) variation, the time-series (within-country) dimension provides some additional information. This information is greatest for variables, such as the terms of trade and inflation, that have varied a good deal over time within countries.

The underlying theory relates to long-term growth, and the precise timing between growth and its determinants is not well specified at the high frequencies characteristic of "business cycles." For example, relationships at the annual frequency would likely be dominated by mistiming and, hence, effectively by measurement error. In addition, many of the variables considered — such as fertility rates, life expectancy, and educational attainment — are not actually measured for many countries at periods finer than 5 or 10 years. These considerations suggest a focus on the determination of growth rates over fairly long intervals. As a compromise with the quest for additional information, I settled on periods of five or ten years; specifically, growth rates were considered for 1965-75 and 1975-85 and for a final five-year period, 1985-90. When the data through 1995 become available, the third period will be lengthened to 1985-95.

The estimation uses an instrumental-variable technique, where some of the instruments are earlier values of the regressors. (The method is three-stage least squares, except that each equation contains a different set of instruments; see the notes to Table 1 for details.) This approach may be satisfactory because the residuals from the growth-rate equations are essentially uncorrelated across the periods. In any event, the regressions describe the relation between growth rates and prior values of the explanatory variables.

[3] The data and detailed definitions of the variables are contained in the Barro-Lee data set, which is available via anonymous FTP from the National Bureau of Economic Research. Updated figures on educational attainment are available from the World Bank web site. An updated version of the full data base will soon be available from this site.

[4] Most of the GDP figures are from version 5.6 of the Summers-Heston data set (see Summers and Heston (1991, 1993) for general descriptions). World Bank figures on real GDP growth rates (based on domestic accounts only) are used for 1985-90 when the Summers-Heston figures are unavailable.

The regression shown in column 1 includes explanatory variables that can be interpreted as initial values of state variables or as choice and environmental variables. The state variables include the initial level of GDP and measures of human capital in the forms of schooling and health. The GDP level reflects endowments of physical capital and natural resources (and also depends on effort and the unobserved level of technology). The choice and environmental variables are the fertility rate, government con-

TABLE 1.

Regressions for Per Capita Growth Rate

independent variable	(1)	(2)
log(GDP)	−0.0254	−0.0225
	(0.0031)	(0.0032)
male secondary and higher schooling	0.0118	0.0098
	(0.0025)	(0.0025)
log(life expectancy)	0.0423	0.0418
	(0.0137)	(0.0139)
log(GDP) *male schooling	−0.0062	−0.0052
	(0.0017)	(0.0017)
log(fertility rate)	−0.0161	−0.0135
	(0.0053)	(0.0053)
government consumption ratio	−0.136	−0.115
	(0.026)	(0.027)
rule-of-law index	0.0293	0.0262
	(0.0054)	(0.0055)
terms-of-trade change	0.137	0.127
	(0.030)	(0.030)
democracy index	0.090*	0.094
	(0.027)	(0.027)
democracy index squared	−0.088	−0.091
	(0.024)	(0.024)
inflation rate	−0.043	−0.039
	(0.008)	(0.008)
Sub Saharan Africa dummy		−0.0042**
		(0.0043)
Latin America dummy		−0.0054
		(0.0032)
East Asia dummy		0.0050
		(0.0041)
R^2	0.58, 0.52, 0.42	0.60, 0.52, 0.47
number of observations	80, 87, 84	80, 84, 87

TABLE 1—*Continued*

* p-value for joint significance of two democracy variables is 0.0006 in column 1 and 0.0004 in column 2.
** p-value for joint significance of three dummy variables is 0.11.

Notes to Table 1:
The system has three equations, where the dependent variables are the growth rate of real per capita GDP for 1965-75, 1975-85, and 1985-90. The variables GDP (real per capita gross domestic product) and male schooling (years of attainment for the population aged 25 and over at the secondary and higher levels) refer to 1965, 1975, and 1985. Life expectancy at birth is for 1960-64, 1970-74, and. 1980-84. The variable log(GDP)*male schooling is the product of log(GDP) (expressed as a deviation from the sample mean) and the male upper- level schooling variable (also expressed as a deviation from the Sample mean). The rule-of-law index applies to the early 1980s (one observation for each country). The terms-of-trade variable is the growth rate over each period of the ratio of export to import prices. The inflation rate is the growth rate over each period of a consumer price index (or of the GDP deflator in a few cases). The other variables are measured as averages over each period. These variables are the log of the total fertility rate, the ratio of government consumption (exclusive of defense and education) to GDP, and the democracy index. Column 2 includes dummy variables for Sub Saharan Africa, Latin America, and East Asia. Individual constants (not shom1) are also estimated for each period.
Estimation is by three-stage least squares (with different instrumental variables used for each equation). The instruments include the five-year earlier value of log(GDP) (for example, for 1960 in the 1965-75 equation); the actual values of the schooling, life-expectancy, rule-of-law, and terms-of-trade variables; and, in column 2, the three area dummy variables.
Additional instruments are earlier values of the other variables except the inflation rate. For example, the 1961)-75 equation uses the averages of the fertility rate and the government- spending ratio for 1960-64. Dummies for former colonies of Spain or Portugal and for former colonies of other countries aside from Britain and France are also included as instruments. (These variables have substantial explanatory power for inflation.) The instrument list also includes the cross product of the lagged value of log(GDP) (expressed as a deviation from the sample mean) with the male schooling variable (expressed as a deviation from the sample mean).
The estimation weights countries equally but allows for different error variances in each period and for correlation of these errors over time. The estimated correlation of the errors for column 1 is −0.13 between the 1965-75 and 1975-85 equations, 0.05 between the 1965-75 and 1985-90 equations, and 0.04 between the 1975-85 and 1985-90 equations. The pattern is similar for column 2. The estimates are virtually the same if the errors are assumed to be independent over the time periods. Standard errors of the coefficient estimates are shown in parentheses. The values and numbers of observations apply to each period individually.

sumption spending, an index of the maintenance of the rule of law, the change in the terms of trade, an index of democracy (political rights), and the inflation rate.

1. Initial Level of GDP

For given values of the other explanatory variables, the neoclassical model predicts a negative coefficient on initial GDP, which enters in the system in logarithmic form.[5] The coefficient on the log of initial GDP has the interpretation of a conditional rate of convergence. If the other explanatory variables are held constant, then the economy tends to approach its long-run position at the rate indicated by the magnitude of the coefficient.[6]

[5]The variable log(GDP) in Table 1 refers to 1965 in the first period, 1975 in the second period, and 1985 in the third period. Five-year earlier values of log(GDP) are used as instruments. The use of these instruments lessens the estimation problems associated with temporary measurement error in GDP.

[6]A full treatment of convergence would also require an analysis of how the various explanatory variables — especially schooling, health, and fertility — respond to the

The estimated coefficient of -0.025 (s.e. $= 0.003$) is highly significant and implies a conditional rate of convergence of 2.5% per year. The rate of convergence is slow in the sense that it would take the economy 27 years to get half way toward the steady-state level of output and 89 years to get 90% of the way. Similarly slow rates of convergence have been found for regional data, such as the U.S. states, Canadian provinces, Japanese prefectures, and regions of the main western European countries (see Barro and Sala-i-Martin (1995a, Ch. 11)).

Figure 2 shows the partial relation between growth and the starting level of GDP, as implied by the regression from column 1 of Table 1. The horizontal axis plots log(GDP) for 1965, 1975, and 1985 for the observations included in the regression sample. The vertical axis shows the corresponding growth rate of GDP after filtering out the parts explained by all explanatory variables other than log(GDP).[7] Thus, the negative slope shows the conditional convergence relation; that is, the effect of log(GDP) on the growth rate for given values of the other independent variables. In contrast to the lack of a simple correlation in Figure 1, the conditional convergence relation in Figure 2 is clearly defined in the graph. Also, the graph indicates that the relation is not driven by a few outliers and does not appear to be nonlinear.

2. Initial Level of Schooling

Education appears in two variables in the system: average years of attainment for males aged 25 and over in secondary and higher schools at the start of each period and an interaction between the log of initial GD P and the years of male secondary and higher schooling. The data on years of schooling are updated and improved versions of the figures reported in Barro and Lee (1993) (and are available from the World Bank web site).

The results show a significantly positive effect on growth from the years of schooling at the secondary and higher level for males aged 25 and over (0.0118 [0.0025]).[8] On impact, an extra year of male upper-level schooling is therefore estimated to raise the growth rate by a substantial 1.2 percentage points per year. (In 1990, the mean of the schooling variable was 1.9 years with a standard deviation of 1.3 years.) The partial relation between the growth rate and the schooling variable — constructed analogously to the method described before for log(GDP) — is shown in Figure 3.

---

development of the economy. Future research will be directed at quantifying these relationships.

[7]The residual is calculated from the regression system that contains all of the variables, including the log of initial GDP. But the contribution from initial GDP is left out to compute the variable on the vertical axis in the scatter diagram. The residual has also been normalized to have a zero mean. The fitted straight line shown in the figure comes from an ordinary-least-squares (OLS) regression of the residual on the log of initial GDP.

[8]Schooling of those aged 25 and over has somewhat more explanatory power than schooling of those aged 15 and over.

FIG. 2. Growth Rate versus Level of GDP

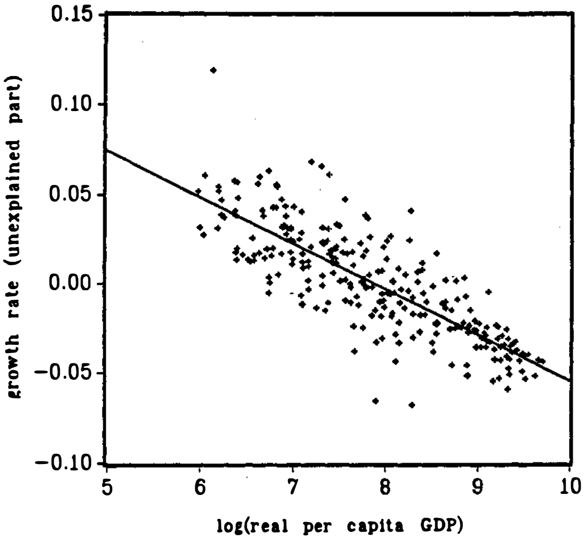
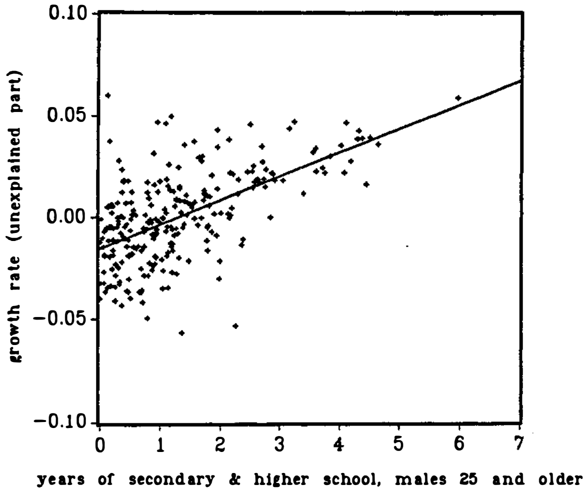


FIG. 3. Growth Rate versus Male Schooling



Male primary schooling (of persons aged 25 and over) has an insignificant effect if it is added to the system; the estimated coefficient is −0.0005 (0.0011), whereas that on upper-level schooling remains similar to that found before (0.0119 [0.0025]). Thus, growth is predicted by male schooling at the upper levels but not by male schooling at the primary level. However, primary schooling is indirectly growth enhancing because it is a prerequisite for training at the secondary and higher levels.

More surprisingly, female education at various levels is not significantly related to subsequent growth. For example, if years of schooling at the secondary and higher levels for females aged 25 and over is added to the system shown in column 1 of Table 1, then the estimated coefficient of this variable is −0.0023 (0.0046), whereas that for males remains significantly positive, 0.0132 (0.0036). For primary schooling of women aged 25 and over, the estimated coefficient is −0.0001 (0.0012), whereas that for men (25 and over for secondary and higher schools) is 0.0118 (0.0025). Thus, these findings do not support the hypothesis that the education of women is a key to economic growth.

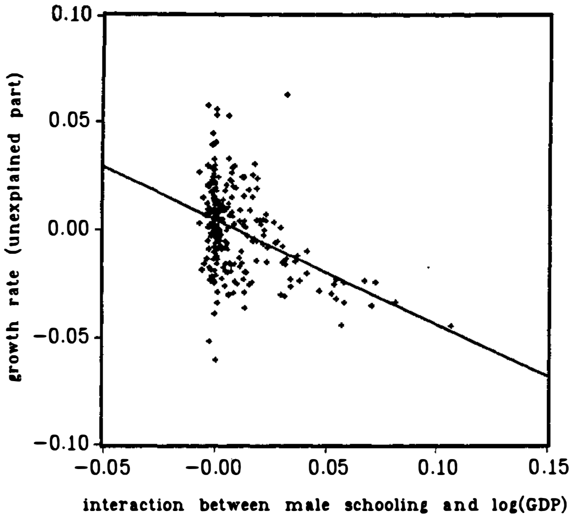
Some additional results indicate that female schooling is important for other indicators of economic development, such as fertility, infant mortality, and political freedom. Specifically, female primary education has a strong negative relation with the fertility rate (see Schultz (1989), Behrman (1990), and Barro and Lee (1994)). A reasonable inference from this relation is that female education would spur economic growth indirectly by lowering fertility, and this effect is not captured in the regressions shown in Table 1 because the fertility rate is already held constant. If the fertility rate is omitted from the system, then the estimated coefficient on female primary schooling (the level of female schooling that affects fertility inversely) is 0.0012 (0.0012), which is positive but not significantly different from zero. Thus, there is only slight evidence that female education enhances economic growth through this indirect channel.

Returning to column 1 of Table 1, the significantly negative estimated coefficient of the interaction term between male schooling and log(GDP), −0.0062 (0.0017), implies that more years of school raise the sensitivity of growth to the starting level of GDP. Starting from a position at the sample mean, an extra year of male upper-level schooling is estimated to raise the magnitude of the convergence coefficient from 0.026 to 0.032. This result supports theories that stress the positive effect of education on an economy's ability to absorb new technologies. The partial relation between the growth rate and the interaction variable appears in Figure 4. (The points at the far right of the diagram are for the most developed countries — such as the United States, Canada, and Sweden-which have high values of GDP and schooling.)

3. Initial Health Status

The population's overall health status is measured here by the log of life expectancy at birth at the start of each period. The results are, however, similar with some alternative aggregate indicators of health, such as the infant mortality rate, the mortality rate up to age five, or life expectancy at age five.

FIG. 4. Growth Rate versus Interaction between Schooling and Level of GDP

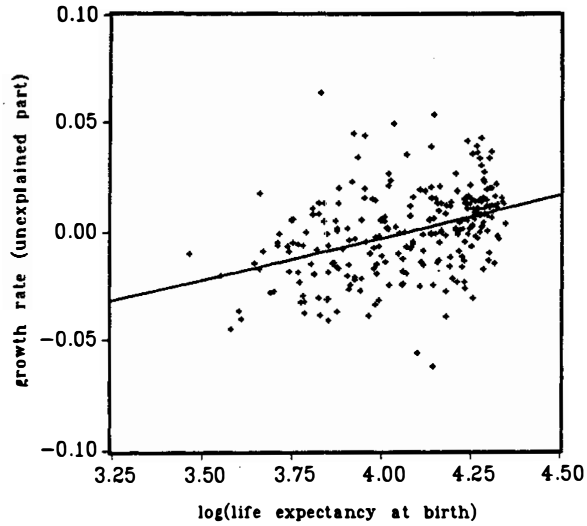


The regression in column 1 reveals a significantly positive effect on growth from initial human capital in the form of health. The coefficient on the log of life expectancy at birth is 0.042 (0.014). This result implies, other things equal, that a rise in life expectancy from 50 to 70 years (that is, by 40 percent) would raise the growth rate on impact by 1.4 percentage points per year. Hence, the link between overall health status and subsequent economic growth appears to be substantial. Moreover, this effect arises even though school attainment and GDP are also included in the regressions.

The partial relation between growth and life expectancy is shown in Figure 5. This figure demonstrates that the relation between health status and subsequent growth is clearly positive, roughly linear (in the log of life expectancy), and is not driven by outliers. An important objective of future research is to extend this finding to more precise indicators of health status, especially those that include the adverse effects of disease. This empirical work will be a central part of the future research in this area.

The World Bank's 1993 Development Report (pp. 25-29) made an interesting attempt to incorporate disease along with mortality through the concept of disease-adjusted life years (DALYs). Lost years of life are computed by comparing actual age of death with the expectation of life in a low-mortality population. Various categories of disabling diseases were added to premature deaths, using weights between 0 and 1 and taking account of the likely duration of the disability. The method for deriving the weights is unclear; the report says that it is a "severity weight that measured the severity of the disability in comparison with loss of life." DALYs

FIG. 5. Growth Rate versus Life Expectancy



were computed as a present value, using a discount rate of 3% and assuming an inverse-u pattern for the value of a year of life between ages 0 and 90. The methodology for constructing the inverse u-pattern was not explained; the report says that it reflects "a consensus judgment, but other patterns could be used."

There seem to be a number of problems with this procedure. First, the discounting procedure and the inverted u-shape for the value of a year of life suggest an attempt to compute the present value of labor earnings, probably including (as is reasonable) an imputed value for productive activities at home. An appropriate version of this procedure would estimate net labor earnings (including imputations) at each date by subtracting expenditures for maintenance and accumulation of human capital. These expenditures include amounts spent on education, health, nutrition, and so on. In this approach, children and old people would typically have a negative contemporaneous contribution to a family's net labor earnings, and the peak in the present value of these earnings would likely occur just after most of the investment in schooling had occurred. The peak in the World Bank DALYs figure at age 10 (Figure 1.3) then seems implausible, except perhaps for the least developed countries.

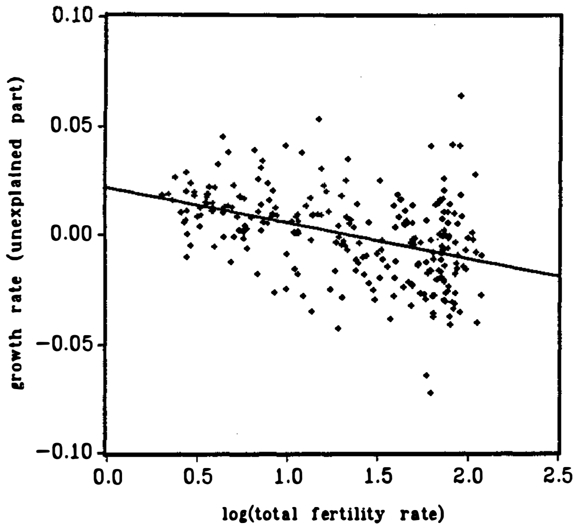
In any event, the present value of net labor earnings does not constitute a reasonable definition of the value of life, as is clear from a consideration of apparently unproductive old people. A better approach is to attempt to estimate amounts that people are willing to pay to accept small increases in the probability of death or disability. Such approaches, exemplified by the research of Rosen (1988), have been used effectively in the U.S. context (and are now frequently used to assess damages in wrongful death cases).

Neither the DALYs approach nor the value-of-life approach connect directly to the link between health and economic growth. In this context, the important aspect of health is its contribution at the margin to current productivity and to the incentives to invest in human capital. These aspects of health are brought out in the model developed in section II of this paper. Further research should be directed at methods to identify the theoretical concepts of health capital and health investment with empirical counterparts.

4. Fertility Rate

If the population is growing, then a portion of the economy's investment is used to provide capital for new workers, rather than to raise capital per worker. For this reason, a higher rate of population growth has a negative effect on y^*, the steady-state level of output per effective worker in the neoclassical growth model. Another, reinforcing, effect is that a higher fertility rate means that increased resources must be devoted to childrearing, rather than to production of goods (see Becker and Barro (1988)). The regression in column 1 shows a significantly negative coefficient, −0.016 (0.005), on the log of the total fertility rate. The partial relation between growth and fertility is in Figure 6.

FIG. 6. Growth Rate versus Fertility Rate



Fertility decisions are surely endogenous; previous research has shown that fertility typically declines with measures of prosperity, especially female primary education and health status (see Schultz (1989), Behrman (1990), and Barro and Lee (1994)). The estimated coefficient of the fertility rate in the growth regression shows the response to higher fertility for given values of male schooling, life expectancy, GDP, and so on. Since the average of the fertility rate over the preceding five years is used as an instrument, the coefficient likely reflects the impact of fertility on growth, rather than vice versa. (In any event, the reverse effect would involve the level of GDP, rather than its growth rate.) Thus, although population. growth cannot be characterized as the most important element in economic progress, the results do suggest that an exogenous drop in birth rates would raise the growth rate of per capita output.

The previous section described a direct positive impact of health status, measured by life expectancy at birth, on economic growth. An additional effect of health on growth would work indirectly through the determination of fertility. In particular, a reduction in mortality rates would likely lower fertility and thereby expand growth as, indicated by the negative coefficient on the fertility rate in the growth regressions. This stimulus to growth would add to the direct effect from improved health.

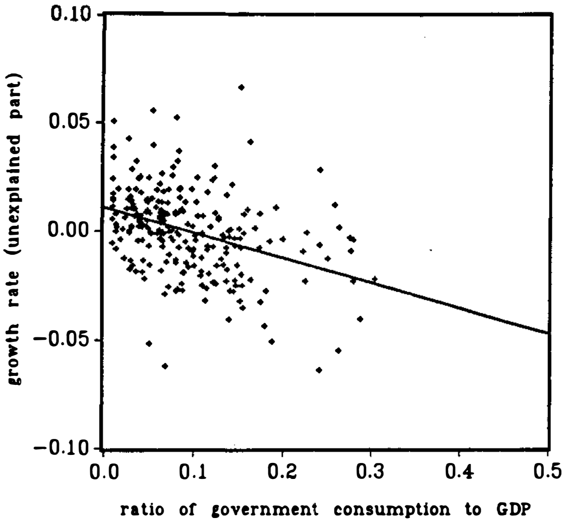
5. Government Consumption

The regression in column 1 of Table 1 also shows a significantly negative effect on growth from the ratio of government consumption (measured exclusive of spending on education and defense) to GDP. The estimated coefficient is −0.136 (0.026). (The period-average of the ratio enters into the regression, and the average of the ratio over the previous five years is used as an instrument.) The particular measure of government spending is intended to approximate the outlays that do not enhance productivity. Hence, the conclusion is that a greater volume of nonproductive government spending — and the associated taxation — reduce the growth rate for a given starting value of GDP. In this sense, big government is bad for growth. The partial relation between growth and the government consumption variable appears in Figure 7.

6. The Rule-of-Law Index

Knack and Keefer (1995) discuss a variety of subjective country indexes prepared for fee-paying international investors by International Country Risk Guide. The concepts covered include quality of the bureaucracy, political corruption, likelihood of government repudiation of contracts, risk of government expropriation, and overall maintenance of the rule of law. (The various time series cover 1982 to 1995 and are available from Political Risk Services of Syracuse, New York.) The general idea is to gauge the attractiveness of a country's investment climate by considering the effectiveness of law enforcement, the sanctity of contracts, and the state of other influences on the security of property rights. Although these data are subjective, they have the virtue of being prepared contemporaneously

FIG. 7. Growth Rate versus Government Consumption Ratio



by local experts. Moreover, the willingness of customers to pay substantial fees for this information is perhaps some testament to their validity.

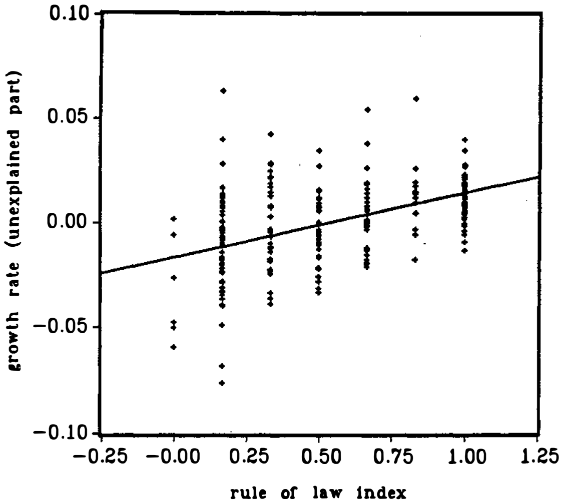
Among the various series available, the indicator for overall maintenance of the rule of law seemed a priori to be most relevant for investment and growth. This indicator was initially measured in 7 categories on a 0 to 6 scale, with 6 the most favorable. The scale has been revised here to 0 to 1, with 0 indicating the worst maintenance of the rule of law and 1 the best.

The rule-of-law variable (observed, because of lack of earlier data, only once for each country in the early 1980s) was included in the regression system reported in column 1 of Table 1 and has a significantly positive coefficient, 0.0293 (0.0054). (The other measures of investment risk, including political corruption, and various indicators of political instability are insignificant in these kinds of growth regressions if the rule-of-law index is also included.) The interpretation is that greater maintenance of the rule of law is favorable to growth. Specifically, an improvement by one rank in the underlying index (corresponding to a rise by 0.167 in the rule-of-law variable) is estimated to raise the growth rate on impact by 0.5 percentage points. The partial relation between growth and the rule-of-law index is in Figure 8. (Note that only seven values for the index are observed.)

7. Democracy

The measure of democracy used in the present study is the indicator of political rights compiled by Gastil and his followers (1982-83 and subsequent issues) from 1972 to 1995. A related variable from Bollen (1990) is used for 1960 and 1965. The Gastil concept of political rights is indicated

FIG. 8. Growth Rate versus Rule of Law Index



by Gastil's basic definition: "Political rights are rights to participate meaningfully in the political process. In a democracy this means the right of all adults to vote and compete for public office, and for elected representatives to have a decisive vote on public policies." (Gastil, 1986-87 edition, p. 7.) In addition to the basic definition, the classification scheme rates countries (somewhat impressionistically) as less democratic if minority parties have little influence on policy.

Gastil applied the concept of political rights on a subjective basis to classify countries annually into 7 categories, where group 1 is the highest level of political rights and group 7 is the lowest. The classification is made by Gastil and his associates based on an array of published and unpublished information about each country. Unlike the rule-of-law index, which was discussed above, the subjective ranking is not made directly by local observers.

The original ranking from 1 to 7 has been converted here to a scale from 0 to 1, where 0 corresponds to the fewest political rights (Gastil's rank 7) and 1 to the most political rights (Gastil's rank 1). The scale from 0 to 1 corresponds to the system used by Bollen.

The system shown in column 1 of Table 1 allows for a quadratic in the democracy indicator. In this case, the estimated coefficients on democracy and its square are each statistically significant. (The p-value for joint significance of the two terms is 0.001.) The pattern of results — a positive coefficient on the linear term and a negative coefficient on the square-means that growth is increasing in democracy at low levels of democracy, but the relation turns negative once a moderate amount of political freedom has been attained. The estimated turning point occurs at an indicator value of approximately 0.5, which corresponds to the levels of democracy in 1995 for Malaysia and Mexico.

One way to interpret the results is that, in the worst dictatorships, an increase in political rights tends to enhance growth and investment because the benefit from limitations on governmental power is the key matter. But in places that have already achieved a moderate amount of democracy, a further increase in political rights impairs growth and investment because the dominant effect comes from the intensified concern with income redistribution.

FIG. 9. Growth Rate versus Indicator of Democracy

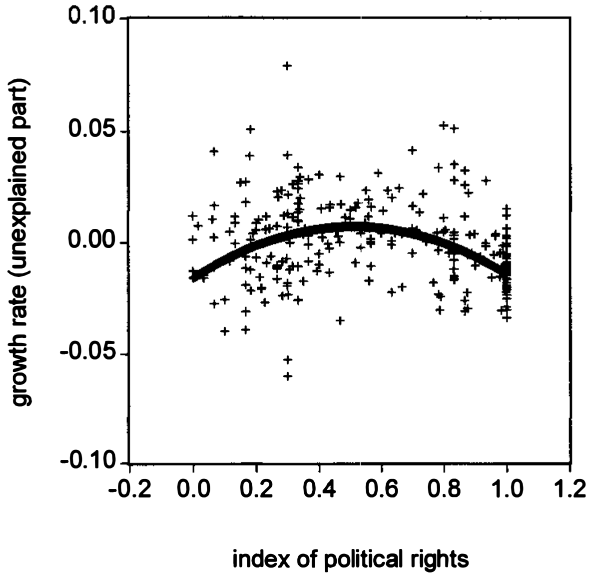


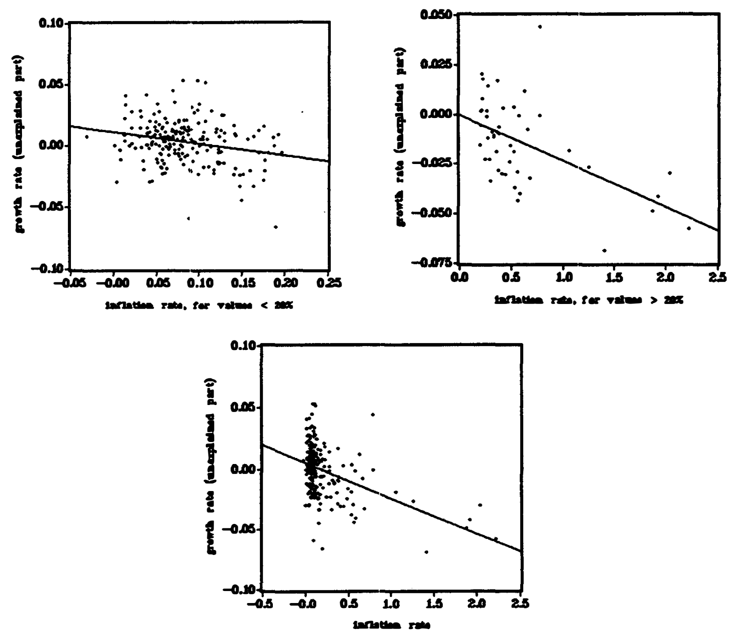
Figure 9 shows the partial relation between the growth rate and the democracy indicator, as implied by the system shown in column 1 of Table 1. (The concentration of points at a democracy value of 1.0 corresponds to the many OECD countries that are rated as fully democratic.) An inverse u-shape can be discerned in the plot, with many of the low and high democracy places exhibiting negative residuals.

The overall relation between growth and democracy is far from perfect; for example, a number of countries with little democracy have large positive residuals. Also, the places with middle levels of democracy seem to avoid low growth rates but not to have especially high growth rates. Thus, there is only the suggestion of a nonlinear relation in which more democracy raises growth when political freedoms are weak but depresses growth when a moderate amount of freedom is already established. One cannot conclude from this evidence that more or less democracy is a critical element for economic growth.

8. Inflation

A key difficulty in isolating the effect of inflation on growth is the endogeneity of inflation; specifically, inflation may react to growth or other aspects of economic performance. For this reason, the estimated effects of inflation shown in Table 1 are based on instrumental variables, where the key instruments in this context are measures of prior colonial status.[9] This procedure exploits the observation that past colonies of Spain, Portugal, and some other countries are much more likely to pursue high-inflation monetary policies than are non-colonies or countries that were previously possessions of France or Britain. For further discussion of these ideas, see Barro (1996, part III).

FIG. 10. Growth Rate versus Inflation Rate



The estimated coefficient of inflation in Table 1 is −0.043 (0.008), which implies that a rise in average annual inflation by 10 percentage points

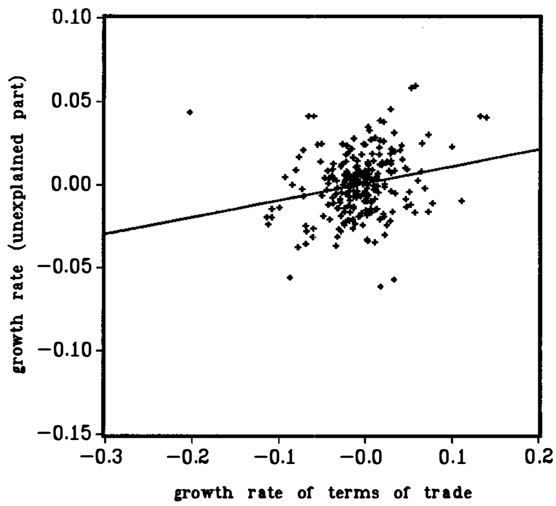
[9]Results are similar, but somewhat reduced in magnitude, if lagged inflation is used as an instrument.

would lower the growth rate by 0.4 percentage points per year. Figure 10 shows the partial relation between growth and inflation for three ranges of inflation: less than 20%, more than 20%, and the entire range. These diagrams show that the main evidence for an adverse effect of inflation on economic growth comes from the experiences of high inflation. For rates of inflation below 10-15% per year, there is not much indication of a systematic relation between growth and inflation.

9. Terms of Trade

Changes in the terms of trade have often been stressed as important influences on developing countries, which typically specialize their exports in a few primary products. The effect of a change in the terms of trade — measured as the ratio of export to import prices — on GDP is, however, not mechanical. If the physical quantities of goods produced domestically do not change, then an improvement in the terms of trade raises real domestic income and probably consumption, but would not affect real GDP. Movements in real GDP occur only if the shift in the terms of trade stimulates a change in domestic employment and output. For example, an oil-importing country might react to an increase in the relative price of oil by cutting back on its employment and production.

FIG. 11. Growth Rate versus Change in Terms of Trade



The result in column 1 of Table 1 shows a significantly positive coefficient on the terms of trade: 0.14 (0.03). (The change in the terms of trade is regarded as exogenous to an individual country's growth rate and is therefore included as an instrument.) Thus, an improvement in the terms of trade apparently does stimulate an expansion of domestic output. The partial relation with growth appears in Figure 11. Although the terms-of-trade variable is statistically significant, it turns out not to be the key element in the weak growth performance of many poor countries, such as those in Sub Saharan Africa.

10. Regional Variables

It has often been observed that recent rates of economic growth have been surprisingly low in Sub Saharan Africa and Latin America and surprisingly high in East Asia. For 1975-85, the mean per capita growth rate for all 124 countries with data was 1.0%, compared with −0.3% in 43 Sub Saharan African countries, −0.1% in 24 Latin American countries, and 3.7% in 12 East Asian countries. For 1985-90, the average growth rate was again 1.0% (for 129 places), compared with 0.1% in 40 Sub Saharan African countries, 0.4% in 29 Latin American countries, and 4.0% in 15 East Asian countries. An important question is whether these regions continue to look like outliers once the explanatory variables considered in Table 1 have been taken into account.

In some previous cross-country regression studies, such as Barro (1991), dummy variables for Sub Saharan Africa and Latin America were found to enter negatively and significantly into growth regressions. However, column 2 of Table 1 shows in the present specification that dummies for these two areas and also for East Asia are individually insignificant. (The p-value for joint significance of the three dummy variables is 0.11.) Thus, the unusual growth experiences of these three regions is mostly accounted for by the explanatory variables.

The inclusion of the inflation rate is critical for eliminating the significance of the Latin America dummy. The Latin America dummy also becomes significant if the fertility rate or the government consumption ratio is omitted. In the case of Sub Saharan Africa, the government consumption ratio is the only individual variable whose omission causes the dummy to become significant. For East Asia, the dummy is significant if male schooling, the rule-of-law indicator, or the democracy variables are deleted.

11. Investment Ratio

In the neoclassical growth model for a closed economy, the saving rate is exogenous and equal to the ratio of investment to output. A higher saving rate raises the steady-state level of output per effective worker and thereby raises the growth rate for a given starting value of GDP. Some empirical studies of cross-country growth have also reported an important positive role for the investment ratio; see, for example, DeLong and Summers (1991) and Mankiw, Romer, and Well (1992).

Reverse causation is, however, likely to be important here. A positive coefficient on the contemporaneous investment ratio in a growth regression may reflect the positive relation between growth opportunities and investment, rather than the positive effect of an exogenously higher investment ratio on the growth rate. This reverse effect is especially likely to apply for open economies. Even if cross-country differences in saving ratios are exogenous with respect to growth, the decision to invest domestically, rather than abroad, would reflect the domestic prospects for returns on investment, which would relate to the domestic opportunities for growth.

The system from column 1 of Table 1 has been expanded to include the period average investment ratio as an explanatory variable. If the instrument list includes the investment ratio over the previous five years, but not the contemporaneous value, then the estimated coefficient on the investment variable is positive, but not statistically significant, 0.027 (0.021). In contrast, the estimated coefficient is almost twice as high and statistically significant if the contemporaneous investment ratio is included as an instrument, 0.043 (0.018). These findings suggest that much of the positive estimated effect of the investment ratio on growth in typical cross-country regressions reflects the reverse relation between growth prospects and investment. Blomstrom, Lipsey, and Zejan (1993) reach similar conclusions in their study of investment and growth.

3. A NEW MODEL OF HEALTH AND ECONOMIC GROWTH

As mentioned before, previous theoretical work on growth has often stressed the role of education as a contributor to human capital but has tended to neglect the role of health. This section describes a framework — effectively an extension of the neoclassical growth model — to incorporate a concept of health capital. A key feature of this analysis is the two-way causation between health and the economy. Better health tends in various ways to enhance economic growth. At the same time, economic advance encourages further accumulation of health capital.

The model includes a direct impact of health on productivity. That is, for given quantities of labor hours, physical capital, and worker schooling and experience, an improvement in health raises a worker's productivity. In addition to this direct effect, an improvement in health lowers rates of mortality and disease and thereby decreases the effective rate of depreciation on human capital; that is, on schooling and health itself. Through this channel, an increase in health raises the demand for human capital and thereby has a further, indirect positive effect on productivity.

In the initial setting, health is viewed as a purely private good that is financed privately. In this context, one can think of health expenditures as involving paid visits to a doctor, unsubsidized purchases of medicine, time and money spent on exercise and nutrition, and so on. Subsequent sections allow for public financing and for spillover effects or externalities that make health a public good.

In the first setting, population growth is treated as exogenous. However, a later section brings in fertility choice. This extension allows for an effect of health capital on fertility and, hence, for an additional influence on economic growth.

3.1. A Basic Model with Health as a Private Good

The output of goods, Y, depends on inputs of physical capital, K, worker schooling (and other aspects of training and experience), S, worker health capital, H, and the amount of labor hours, L. To simplify matters, we assume that production takes the Cobb-Douglas form,

$$Y = A \cdot K^{\alpha} S^{\beta} H^{\gamma} (Le^{xt})^{1-\alpha-\beta-\gamma}, \tag{2}$$

Where $\alpha > 0, \beta > 0, \gamma > 0$, and $0 < \alpha + \beta + \gamma < 1$. The formulation therefore assumes constant returns to scale in the four inputs and diminishing returns with respect to each input individually. The parameter $A > 0$ is the exogenous baseline level of technology, and $x \geq 0$ is the exogenous rate of labor-augmenting technological progress. The total of labor input, L, is assumed to correspond to population, so that variations in the ratio of work effort to population are not considered.

The idea in equation (2) is that output depends not only the conventional inputs — physical capital, raw labor, and human capital in the form of schooling — but also on the state of worker health, H. This health capital could influence worker energy, effort, reliability, and so on.

It is convenient to divide through by the quantity of effective labor input, Le^{xt}, On both sides of equation (2) to express the production function in intensive form,

$$\hat{y} = A \cdot \hat{k}^{\alpha} \cdot \hat{s}^{\beta} \cdot \hat{h}^{\gamma}, \tag{3}$$

Where $\hat{y} \equiv Y/Le^{xt}$, $\hat{k} \equiv K/Le^{xt}$, $\hat{s} \equiv S/Le^{xt}$, and $\hat{h} \equiv H/Le^{xt}$ are quantities per unit of effective labor. Amounts of output and capital per Unit of labor (or per person) will be denoted correspondingly by y, k, s, and h.

The representative household-producer in the economy is assumed to maximize utility over an infinite horizon, as given by the standard form,[10]

$$U = \int_0^{\infty} \left(\frac{c^{1-\theta} - 1}{1 - \theta} \right) \cdot e^{nt} e^{-\rho t} dt, \tag{4}$$

where c is consumption per person, $\rho > 0$ is the constant rate of time preference, and $\theta > 0$ is the constant elasticity of marginal utility (with $\theta = 1$ corresponding to log utility). A lower value of the parameter θ signifies

[10]See, for example, Barro and Sala-i-Martin (1995, Ch. 2). The model can be worked out equivalently with households separated from firms.

that the household is more willing to substitute consumption over time. The parameter $n \geq 0$ is the exogenous and constant rate of population growth within the household. The model in the next section allows for an endogenous determination of fertility and, hence, population growth. The infinite horizon assumed in equation (4) is a convenient device that can be thought of in terms of a family dynasty that goes on indefinitely through altruistic linkages between parents and children.

Each household-producer has access to the production technology shown in equation (3) and chooses the paths of consumption, c, and the amounts to invest in the three types of capital to maximize utility in equation (4). The household starts at the arbitrary initial date 0 with given endowments of the three types of capital. At each point in time, output can be divided among consumption and gross investment in the three kinds of capital. Units of consumption and each type of investment substitute on a fixed basis in accordance with the usual one-sector production model. Thus, the assumption is that the production of consumables, physical capital, schooling, and health all involve the same factor intensities.[11]

The gross investment flows per unit of effective labor — denoted by $\hat{i}_k$, $\hat{i}_s$, and $\hat{i}_h$ — determine the evolution of capital stocks as follows:

$$\dot{\hat{k}} = \hat{i}_k - (\delta + x + n) \cdot \hat{k}, \quad (5)$$

$$\dot{\hat{s}} = \hat{i}_s - (d + x + n) \cdot \hat{s}, \quad (6)$$

$$\dot{\hat{h}} = \hat{i}_h - (d + x + n) \cdot \hat{h}, \quad (7)$$

where a dot over a variable indicates differentiation with respect to time, $\delta > 0$ is the exogenous depreciation rate for physical capital, and $d > 0$ is the depreciation rate for schooling and health. The household's budget constraint is

$$\hat{y} = \hat{c} + \hat{i}_k + \hat{i}_s + \hat{i}_h. \quad (8)$$

A key assumption is that the depreciation rate, d, for human capital is a decreasing function of the stock of health capital per person, h:

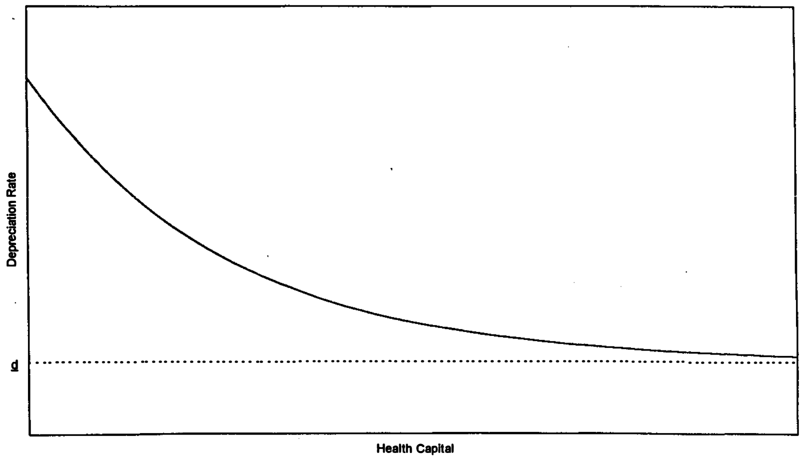
$$d = \underset{(-)}{d(h)} \quad (9)$$

One effect here is that better health reduces the probability of death. If a family member dies, then the dynasty loses the schooling and other human capital possessed by that member; in this sense, a higher mortality rate

[11] Barro and Sala-i-Martin(1995, Ch. 5) consider a two-sector production model for goods and human capital in the form of education. In this setting, the production of human capital is assumed to be relatively intensive in human capital.

amounts to a higher depreciation rate. The variable d in equations (6) and (7) captures this mechanism (although, for convenience, in a deterministic form). This kind of effect would also apply to non-fatal diseases that effectively deteriorate a person's human capital. Thus, empirical implementations of the model should include the burden of disease in addition to the adverse effects of mortality.

FIG. 12. Human Capital Depreciation versus Health Capital



We assume that the derivative in equation (9) is negative, but that the magnitude of this derivative decreases as h rises. Hence, the relation between d and h takes the form shown in Figure 12. The additional assumption in the figure is that d has a positive lower bound, d_0, so that no amount of health capital can lower the depreciation rate below this value. This pattern would emerge if life expectancy has a finite upper bound.

Physical capital differs from human capital because a person's death or disease does not directly affect machines and buildings. Hence, the depreciation rate δ in equation (5) does not depend on the state of health. This independence of depreciation from health would also apply to some forms of social human capital; for example, to knowledge about diseases and medicines or to technological advances more broadly.

The household's dynamic optimization program is a standard problem, which can be solved by familiar methods.[12] One of the resulting first-order conditions describes the choice of consumption over time. The two other conditions dictate equality among the rates of return on the three types of capital at all points in time. If we assume that the constraints

[12]These methods are described in Barro and Sala-i-Martin (1995, Ch.2 and appendix).

of nonnegative gross investment in the three types of capital are never binding, then the solutions are interior and fairly simple to write down.

The evolution of consumption over time is given by a familiar condition:

$$g_c \equiv \dot{c}/c = (1/\theta) \cdot \left(\alpha A \cdot \hat{k}^{\alpha-1} \cdot \hat{s}^{\beta} \cdot \hat{h}^{\gamma} - \delta - \rho\right) \tag{10}$$

The first term within the parentheses on the right-hand side is the gross marginal product of physical capital. The subtraction of δ from this gross marginal product gives the net marginal product, which must equal the rate of return in the economy, r. Equation (10) says that consumption growth is an increasing function of the difference between r and the rate of time preference, ρ. The larger the household's willingness to substitute consumption over time, as implied by a smaller value of θ, the greater is the sensitivity of consumption growth to the gap between r and ρ. Given the simple form of utility function in equation (4), the condition in equation (10) equates the household's intertemporal rate of substitution for consumption to the rate of return.

The equality among rates of return for the three types of capital can be written as

$$\begin{aligned} \alpha A \cdot \hat{k}^{\alpha-1} \cdot \hat{s}^{\beta} \cdot \hat{h}^{\gamma} - \delta &= \beta A \cdot \hat{k}^{\alpha} \cdot \hat{s}^{\beta-1} \cdot \hat{h}^{\gamma} - d \\ &= \gamma A \cdot \hat{k}^{\alpha} \cdot \hat{s}^{\beta} \cdot \hat{h}^{\gamma-1} - (s+h) \cdot (\partial d/\partial h) - d \end{aligned} \tag{11}$$

The first line is straightforward, as it shows the equality between the net marginal products of physical capital and schooling. One important effect here, which has been stressed in an analogous context by Meltzer (1995), is that an increase in health capital lowers the human-capital depreciation rate, d, and thereby raises the rate of return for investments in schooling. To put it another way, an increase in life expectancy increases the incentive to invest in education.

Meltzer's model also includes a rich structure of mortality by age, so that he can distinguish mortality rates of children from those of adults. He observes that the critical influence on schooling investment involves mortality rates at adult ages in which education affects productivity. In particular, infant and old-age mortality do not have the same interplay with the demand for education. Ehrlich and Lui (1991) also distinguish mortality rates for children from those of adults in terms of the interaction with the demand for human capital.

The second line of equation (11) includes as a component of the rate of return on health capital the negative effect of better health on the human-capital depreciation rate, $\partial d/\partial h$. This influence is more important the larger the capital stock per person, $s + h$, to which the depreciation rate d applies. Note that one aspect of more health is that it lowers d and thereby raises the rate of return on further investment in health. (As an example, control of an array of infectious diseases raises life expectancy and thereby increases the rate of return for investments in the treatment of old-age ailments, such as heart disease and cancer.) However, health investments tend to encounter diminishing returns overall because of the declining direct productivity effect (through the term $\hat{h}^{\gamma-1}$ in the second line of equation (11)) and the tendency for $\partial d/\partial h$ to fall as h rises.

Much of the dynamics of the model is analogous to that of the conventional neoclassical growth model, which has a single form of capital. If the economy begins with small endowments of each type of capital per unit of effective worker, then the A dynamic path involves rising values of $\hat{k}$, $\hat{s}$, and $\hat{h}$.[13] Because of the diminishing returns to the three types of capital overall, as assumed for given raw labor in equation (1), the path also tends to involve diminishing rates of return and growth rates.

If we neglected the effect of health on the human-capital depreciation rate — that is, took $\partial d/\partial h = 0$ — and also assumed that the rates of depreciation on physical and human capital were the same, $\delta = d$, then equation (11) would imply that the ratios of the various capital stocks would remain constant along the dynamic path. In other words, the stocks of physical capital, schooling, and health would always grow at the same (not necessarily constant) rate.

An interesting new property of the present model is that the inverse effect of rising health on the human-capital depreciation rate, d, tends to raise the ratios of schooling and health to physical capital. That is, s/k and h/k tend to rise as the economy develops. The ratios s/y and h/y tend also to increase as the economy becomes richer; that is schooling and health would be relatively more important in higher-income places.

The evolution of the relative quantities of the two types of human capital, h/s, depends on the behavior of the term $(s + h) \cdot \partial d/\partial h$ in equation (11). If this term rises in magnitude over time (because $s + h$ increases), then h/s would tend to rise. However, if the term declines in magnitude over time (because the size of $\partial d/\partial h$ falls), then h/s would tend to decline. We think that the typical pattern will be for a poor country to experience an initial rise in h/s because the dominant effect will come from the increasing importance of human capital, $s+h$. However, the decreasing size of $\partial d/\partial h$, as shown in Figure 12, may eventually dominate and cause h/s to fall at

[13]To get an interior solution at all points in time, we implicitly have to allow the household to rearrange its existing total quantity of capital among the three components. That is, the household would have to be able to exchange schooling capital for physical capital, and vice versa. More realistically, some transitional dynamics would be involved if the household begins with relative stocks of capital that depart greatly from the desired values. Barro and Sala-i-Martin (1995, Ch. 5) provides an example of this kind of dynamics.

a higher range of incomes. We plan to investigate this behavior more fully through dynamic simulations.

In the steady state of the conventional neoclassical model, the quantity of physical capital per unit of effective labor, $\hat{k}$, has risen enough to lower the rate of return on capital to a. value consistent with consumption per person, c, growing at the rate of technological progress, x. Equation (10) implies that this rate of return is given by $\rho + \theta x$. In the steady state, y, k and c all grow forever at the constant rate of technological progress, x.

The analogous situation in the present model is for the quantities of human capital per person, s and h, to grow in the steady state at the rate x, so that $\hat{s}$ and $\hat{h}$ are constant. The problem, however, is that rising h (not $\hat{h}$) implies falling d, which continually raises the rate of return to investments in human capital and is therefore inconsistent with the usual kind of steady state.

If the human-capital depreciation rate, d, behaves as shown in Figure 12, then the model asymptotically approaches a steady state in which d equals its lower bound, d_0, and the quantities $\hat{y}$, $\hat{k}$, $\hat{s}$, and $\hat{h}$ are all constant. In this case, y, k, s, h, and c all grow at the constant rate x. These results are consistent with equation (11) because d is constant and $\partial d/\partial h$ is zero in this steady state.

3.2. Health Services as a Publicly Subsidized Private Good

This section maintains the assumption that health services are a private good; in particular, no spillover effects or externalities are involved. Nevertheless, some portion of these services are assumed to receive a public subsidy or to be provided directly by the government. The first version assumes that the government supplies a designated quantity of health services to each individual and finances this spending by a proportional tax on output. Households are allowed to supplement their health services by buying amounts above the public ration from the private market. The second version assumes that the government establishes a price below marginal cost for medical services and allows people to buy their desired quantity at this price. The subsidy is again financed by a proportional tax on output.

1. A Rationed Quantity of Publicly Provided Medical Services

Let e_h be the quantity of health services provided by the government at a point in time to each person. Although these goods are publicly provided, they are private (that is, rival) in the sense of being useful only to the individual to whom they are supplied. Thus, e_h might represent a doctor's services provided free of charge by the government to an individual patient.

The evolution of a household's health capital from equation (7) is modified to

$$\dot{\hat{h}} = (\hat{e}_h + \hat{i}_h) - (d + x + n) \cdot \hat{h}, \qquad (12)$$

where $\hat{e}_h$ is the amount of public health spending per unit of effective labor. Thus, the flow $\hat{e}_h$ comes "for free" to each household, and this public provision is supplemented by the flow $\hat{i}_h$ of private spending, where $\hat{i}_h \geq 0$ must hold. The assumption in equation (12) is that public and private spending are perfect substitutes as contributors to the accumulation of a household's health capital. If $\hat{e}_h$ is high enough, then the inequality constraint $\hat{i}_h \geq 0$ will bind, and the household will make zero private expenditures on health.

The government is assumed to pay for its spending on health by a proportionate levy at the rate τ on gross output:

$$\hat{e}_h = \tau \hat{y}, \tag{13}$$

Note that no other form of public expenditure (for example, on education) is contained in this model. Each household takes $\hat{e}_h$ and τ as given and recognizes that it retains only the fraction $1-\tau$ of its gross income to use for consumption or investment. Therefore, the household's budget constraint from equation (8) is modified to

$$(1-\tau)\cdot\hat{y} = \hat{c} + \hat{i}_k + \hat{i}_s + \hat{i}_h. \tag{14}$$

If the inequality restriction $\hat{i}_h \geq 0$ is not binding, then the only modification to the previous analysis is that the gross marginal product of each type of capital is multiplied by $1-\tau$ to determine private, after-tax rates of return. For example, the condition for the growth rate of consumption from equation (10) is modified to

$$g_c \equiv \dot{c}/c = (1/\theta)\cdot[(1-\tau)\cdot\alpha A\cdot\hat{k}^{\alpha-1}\cdot\hat{s}^{\beta}\cdot\hat{h}^{\gamma} - \delta - \rho]. \tag{15}$$

For given values of $\hat{k}$, $\hat{s}$, and $\hat{h}$, a higher τ depresses the after-tax rate of return on the right-hand side and therefore lowers the growth rate of consumption. Similarly, the three gross marginal products — for physical capital, schooling, and health — that appear in equation (11) are each multiplied by $1-\tau$ in the new solution.

In this model, the only effect from the public provision of a designated quantity of health services comes from the distorting tax finance. The tax rate on gross output translates into a reduced private rate of return on all types of investment. The consequence is a lower rate of economic growth during the transition and smaller steady-state values of all types of capital, $\hat{k}$, $\hat{s}$, and $\hat{h}$. Since there were no sources of private market failure in the initial setting — in particular, because all goods including health are private — the public intervention reduces the utility attained by the representative household.

If $\hat{e}_h$ is high enough so that the constraint $\hat{i}_h \geq 0$ binds at some points in time, then there is an additional effect involving changes in relative demands for the various kinds of capital. Health capital, $\hat{h}$, tends to rise in relation to physical capital, $\hat{k}$, and schooling, $\hat{s}$. The negative effect on schooling, $\hat{s}$, is mitigated because the higher level of health lowers the human-capital depreciation rate, d. Therefore, $\hat{s}$ would tend to rise in relation to $\hat{k}$. Since the underlying goods are all private, these shifts represent distortions that tend to depress the attained level of utility.

2. Publicly Subsidized Medical Services

Suppose now that the government sets a subsidized price P_h for privately purchased medical services, where $0 < P_h < 1$.[14] (The social marginal cost of health services is still 1.) The government's budget constraint is now

$$(1 - P_h) \cdot \hat{i}_h = \tau \hat{y}, \tag{16}$$

that is, the government pays for the excess of marginal cost over price on the privately chosen amount of spending $\hat{i}_h$. Public finance still takes the form of a proportionate levy at the rate τ on gross output.

The representative household's budget constraint is now changed from equation (14) to

$$(1 - \tau) \cdot \hat{y} = \hat{c} + \hat{i}_k + \hat{i}_s + P_h \hat{i}_h. \tag{17}$$

The new element is the presence of the subsidized price P_h on the right-hand side.

In the solution to this model, the term $1-\tau$ multiplies the gross marginal products of each type of capital in equations (10) and (11), just as in the previous version. The new consideration is that the private rate of return to health investment also takes account of the public subsidy. Therefore, the expression for the rate of return to private health spending in equation (11) becomes

$$\left[\frac{(1-\tau)}{P_h}\right] \cdot \gamma A \cdot \hat{k}^{\alpha} \cdot \hat{s}^{\beta} \cdot \hat{h}^{\gamma - 1} - (s + h) \cdot (\partial d / \partial h) - d.$$

Hence, $P_h < 1$ offsets the effect from a positive tax rate, τ.

The effect of the tax distortion, τ, is to reduce the incentive to invest in all types of capital, just as in the previous version. The price subsidy, $P_h < 1$, induces a substitution in favor of health capital. As in the model in which the rationed quantity of public health services was high enough

[14] The private demand for medical services would be infinite in this model if $P_h = 0$. In practice, however, the private price would include the time and irritation required to receive medical services, so that a finite amount would be demanded even if (as in the present model) the marginal product of health services were always positive.

for the inequality restriction on private spending, $\hat{i}_h \geq 0$, to bind, this intervention tends to raise $\hat{h}$ relative to $\hat{k}$ and $\hat{s}$. The reduction in the human-capital depreciation rate, d, tends again to increase $\hat{s}$ relative to $\hat{k}$. Since the underlying goods are all private, the government's interventions lower attained utility.

3.3. Health Services as a Public Good

A key reason for public provision of health services (aside from redistribution of income or insurance) is that these services actually are a public good. For example, water and air quality and immunization programs involve spillover effects so that each household's utility depends directly on the actions of other households.[15] For a pure public good of this type, the cumulation of a household's health capital (or state of health) depends not on private spending, $\hat{i}_h$, as in equation (7), but on spending in the whole economy. Not surprisingly, the purely private equilibrium will not be Pareto optimal in this circumstance.

Suppose that the cumulation of a household's health capital is given as a modification of equation (7) by

$$\dot{\hat{h}} = \overline{i_h} - (d + x + n) \cdot \hat{h}, \tag{18}$$

where $\overline{i_h}$ is the health spending per unit of effective labor in the overall economy. Since each individual household is small it disregards the contribution of its own health spending to $\overline{i_h}$. Therefore, a purely private equilibrium entails zero spending by all households on health investment. Since the total of each person's health expenditure actually determines $\overline{i_h}$, which matters for each household, there is an externality that leads to underinvestment in health. That is, the socially optimal value of $\overline{i_h}$ — the value that internalizes the spillover effects across households — is positive. This value can be determined by working through a hypothetical "social-planner's" problem for maximizing the representative household's utility. In the present case, the solution is analogous in form to that worked out for the first model with purely private (and, hence, internalized) health spending.

The government can, in principle, choose a level of public health spending per unit A of effective labor, $\hat{e}_h = \overline{i_h}$, to maximize the utility of the rep-

[15] Another type of public good in the health area involves basic research on medical products and procedures. This research requires substantial outlays up front, including the costs1 of failed efforts. However, the production and use of the successful products likely entails only small costs afterwards. Ex-post efficiency requires goods to have low prices, corresponding to their small marginal costs of production, but this procedure fails to provide the necessary reward for research efforts. Hence; a high rate of economic progress may require patent or other protection that secures monopoly profits ex post. A model of this type is worked out in Barro and Barro (1996, part II).

resentative household. This value would coincide with the social-planner's choice of $\overline{i_h}$ if the public spending were financed by a lump-sum (non-distorting) tax. However, if the government must use distorting finance, such as the previously prescribed proportional tax at rate on gross output, then a tradeoff arises. Public determination of the level of health services is desirable to internalize the spillover effects but undesirable because of the adverse effect of the required taxation on after-tax rates of return. Optimization would typically result in levels of public health expenditure that were positive but less than the social planner's values.

From an empirical standpoint, a key issue will be the differential effects on growth from private and public expenditures on health. In particular, some aspects of the spillover effects of health can be assessed by isolating the impact of public health spending on economic growth. The data on public health expenditures for many countries appear to be adequate to carry out this empirical research.

3.4. Health and the Distribution of Income

Caselli and Ventura (1996) show how to extend the neoclassical growth model to incorporate heterogeneity among households. They allow for a distribution of initial wealth, intrinsic productivity, and preferences for smoothing consumption over time. A key finding is that this extension preserves most of the neoclassical model's implications for the behavior of economic growth and other aggregate variables. However, the extended model can be used to assess the evolution of inequality over time and to evaluate the effects of various policies on inequality.

The combination of the Caselli-Ventura framework with the previously described extensions to include health capital would allow for an analysis of the effects of various health policies on income distribution. However, for many purposes, it seems that these effects would be analogous to those that apply to a range of governmental programs, such as welfare payments and educational spending. Obviously, public health programs or provision of "free" health services tend to redistribute resources toward the poor. But a special distinguishing role of health in the context of income distribution would seem to arise only if poor people are unable to make efficients decisions about health and are therefore better off when the government dictates how much of their income be used for health purposes. Surely the government has an important role in disseminating information about disease, sanitation, nutrition, and so on. But it is unclear that the poor are benefited when the government provides them with health services in lieu of cash. It seems that the main case for public health spending comes from the public-goods nature of an array of health activities, not from arguments about the distribution of income. In any event, further research would consider these distributional issues in detail.

3.5. The Interplay between Health and Fertility

The objective here is to study the interaction between health and fertility and thereby to assess further linkages between health and economic growth. The present formulation lays out an analytical framework but does not yet fully assess the workings of this framework. The model is a modification of the treatment in Barro and Sala-i-Martin (1995, Ch. 9), which was an extension of the work in Becker and Barro (1988), Barro and Becker (1989), and Becker, Murphy, and Tamura (1990).

Imagine that a household or family starts at time 0 with 1 member. The number of persons, N, in the family grows due to fertility but falls due to mortality. (This model does not consider marriage and therefore thinks of fertility in an asexual manner.) Gross fertility is modeled for tractability as a flow of persons at the rate n. From this gross flow, a fraction $\iota > 0$ dies immediately — as a way to represent infant mortality — so that $n \cdot (1 - \iota)$ is the net flow into the stock of people.

The mortality rate $d > 0$ applies to all persons, N, who have survived infancy. The rate d reflects adult mortality, but in a simple way that does not consider age, except for the distinction between infants and adults. The change in family size (treated as a continuous flow) is given by

$$\dot{N} = [n \cdot (1 - \iota) - d] \cdot N. \tag{19}$$

In this setting, new persons effectively become adults as soon as they survive infancy.

At present, we treat the two mortality rates, ι and d, as exogenous. The choice of gross fertility rate, n — and, hence, family size, N — will be determined endogenously. Therefore, this model will be able to consider the effects of infant and adult mortality on fertility and economic growth. In future research, the mortality rates and d will be determined endogenously along the lines of the model described in the previous sections. This extended framework will allow for reverse effects from fertility and growth to health and, hence, mortality rates.

The representative household's utility function now includes the size of the family (or number of descendants) at each date, N, in addition to the path of consumption per person, c. We use the functional form,

$$U = \int_0^\infty \left[\frac{(N^\psi c)^{1-\theta} - 1}{1 - \theta} \right] \cdot e^{-\rho t} dt, \tag{20}$$

where $\psi > 0$ and $\psi \cdot (1 - \theta) < 1$. (An analogous form is discussed in Barro and Sala-i-Martin (1995, Ch. 9).) This formulation extends equation (4) to allow for positive and diminishing marginal utility with respect to family size, N. (The term for exogenous population growth, e^{nt}, in equation (4) no longer appears.)

A key part of the analysis is the specification of the costs of bearing and raising children. We assume that this cost is primarily a time cost, with each unit of time valued at the (possibly implicit) wage rate, w. The amount of time required depends on the number of children, which is represented in this continuous-time model by the gross fertility rate, n. Hence, the child-rearing cost takes the form,

$$\text{child-rearing cost} = w \cdot g(n), \tag{21}$$

where the time required, $g(n)$, satisfies $g'(n) > 0$. The marginal cost of child rearing may decrease with n over some range because of fixed costs associated with setting up a family and having children. We assume, however, that $g''(n) > 0$ applies for sufficiently large n. (We could also extend the analysis to include a goods cost of raising children; that is, a cost that does not depend on w.)

The representative household-producer has assets k per person, and these assets earn the rate of return r. All N adult members of the household are assumed to work one unit of time and earn the wage rate w (although labor-force participation and time spent at school or in retirement could also be considered). Therefore, the budget constraint is

$$\dot{k} = w + rk - [n \cdot (1 - \iota) - d] \cdot k - w \cdot g(n) - c. \tag{22}$$

The term $[n \cdot (1 - \iota) - d] \cdot k$ gives the reduction in assets per person, k, due to population growth at the rate $n \cdot (1 - \iota) - d$.

Each household-producer treats the path of prices, r and w, as given and then chooses the path of c and n to maximize utility in equation (20). This maximization is subject to the budget constraint in equation (22), the determination of family size from equation (19), and to a given initial amount of assets per person, $k(0)$.

The first-order conditions can be derived from standard dynamic optimization methods, although the model now involves two state variables, k and N. The analysis is fairly complicated and simplifies only for the case of log utility ($\theta = 1$). We skip all the algebra here and just write down the final results for this case.

The condition for consumption growth is now

$$g_c \equiv \dot{c}/c = r - \rho - [n \cdot (1 - \iota) - d]. \tag{23}$$

To see the parallel with equation (8), substitute the net marginal product of capital for r and set $\theta = 1$. The new element in equation (23) is the subtraction of the rate of population growth from $r - \rho$ on the right-hand side.

The other first-order condition relates the choice of gross fertility, n, to the level of consumption, c:

$$\omega \cdot g'(n) = (1-\iota) \cdot \left[\left(\frac{\psi}{\rho} \right) c - k \right]. \tag{24}$$

This result can be viewed as the equation of the marginal cost of creating an additional surviving person, which is given by $w \cdot g'(n)/(1-\iota)$, to the marginal benefit, which turns out to be given by $(\frac{\psi}{\rho}) \cdot c - k$. If $g''(n) > 0$, then equation (24) implies that n rises with c/w, falls with k/w, falls with the infant mortality rate, ι, and rises with the parameter ψ, which governs the marginal utility of descendants. The adult mortality rate, d, will influence n indirectly through effects on the paths of c and k. Also, aside from its direct negative effect (the opposite to the standard relation), a change in the infant mortality rate, ι, will influence n indirectly by affecting the paths of c and k.

The production function (for example, a Cobb-Douglas form) dictates the relation of w to k. Then the full model, which includes equation (23), determines the paths of c, k, and n. Barro and Sala-i-Martin (1995, Ch. 9) show for an analogous model that the gross fertility rate, n, tends to decline during the process of economic development, as typically found in the data. Also, shifts that tend to reduce n (such as a decline in the taste parameter ψ) tend to spur economic growth along the lines found in the empirical work across countries.

Future research will assess the model's full implications for the effects of the infant and adult mortality rates, ι and d, on fertility and growth. An important part of this work will be the distinction in the predicted effects from reductions in mortality rates at different ages. Also to be considered are alternative specifications that modify the forms of household preferences and child-rearing costs. Later on, the mortality rates will be determined endogenously from the choice of health capital over time.

4. FUTURE RESEARCH PLANS

Planned research will begin with extensions of the theoretical models on health and economic growth. Areas that will receive especial attention include the dynamics of health capital in relation to schooling and physical capital, the interplay between health and fertility, and the effects of health policies on the distribution of income.

The empirical work will start with modifications of previous work on cross–country growth regressions. A central element here is the specification of aggregate measures of health status. The previous reliance on life expectancy at birth will be extended to consider mortality rates at various ages and to allow for the effects of disease. Some disaggregation among specific categories of disease will be considered here. This study will also consider appropriate ways to combine mortality and morbidity to come up with appropriate aggregate indicators of the burden of poor health.

Another area to be examined concerns the distinction between public and private health spending. A key policy question here is to evaluate the rate of return to investment in the public health area.

REFERENCES

Agbion, Philippe and Peter Howitt, 1992. A Model of Growth through Creative Destruction. *Econometrica* **60, 2**, 323-351.

Arrow, Kenneth J., 1962. The Economic Implications of Learning by Doing. *Review of Economic Studies* **29**, 155–173.

Barro, Robert J., 1991. Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics* **106, 2**, 407-433.

Barro, Robert J., 1996. Determinants of Economic Growth: A Cross-Country Empirical Study. National Bureau of Economic Research, working paper no.5698, August.

Barro, Robert J. and Jason R. Barro, 1996. Three Models of Health and Economic Growth. Unpublished, Harvard University, September.

Barro, Robert J. and Gary S. Becker, 1989. Fertility Choice in a Model of Economic Growth. *Econometrica* **57, 2**, 481-501.

Barro, Robert J. and Jong-Wha Lee, 1993. International Comparisons of Educational Attainment. *Journal of Monetary Economics* **32**, 363-394.

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

Becker, Gary S. and Robert J. Barro, 1988. A Reformulation of the Economic Theory of Fertility. *Quarterly Journal of Economics* **103, 1**, 1-25.

Becker, Gary S., Kevin M. Murphy, and Robert Tamura, 1990. Human Capital, Fertility, and Economic Growth. *Journal of Political Economy* **98, 5, part II**, 512-537.

Blömstrom, Magnus, Robert E. Lipsey, and Mario Zejan, 1993. Is Fixed Investment the Key to Economic Growth ? National Bureau of Economic Research working paper no. 4436, August.

Bollen, Kenneth A., 1990. Political Democracy: Conceptual and Measurement Traps. *Studies in Comparative International Development* **Spring**, 7-24.

Caselli, Francesco and Jaume Ventura, 1996. A Representative Consumer Theory of Distribution. Unpublished, Harvard University.

Cass, David, 1965. Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies* **32**, 233-240.

DeLong, J. Bradford and Lawrence H. Summers, 1991. Equipment Investment and Economic Growth. *Quarterly Journal of Economics* **106, 2**, 445-502.

Ehrlich, Isaac and Francis T. Lui, 1991. Intergenerational Trade, Longevity, and Economic Growth. *Journal of Political Economy* **99, 5**, 1029-1059.

Gastil, Raymond D. and followers, 1982. *Freedom in the World, West port CT.* Greenwood Press.

Grossman, Gene M., and Elhanan Helpman, 1991. *Innovation and Grotufh in the Global Economy*. Cambridge MA, MIT Press.

Knack, Stephen and Philip Keefer, 1995. Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures. *Economics and Politics* **7**, 207-227.

Koopmans, Tjalling C., 1965. On the Concept of Optimal Economic Growth, in The Econometric Approach to Development Planning, Amsterdam, North Holland.

Lucas, Robert E., Jr., 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics* **22, 1**, 3-42.

Mankiw, N. Gregory, David Romer, and David N. Well, 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* **107, 2**, 407-437.

Meltzer, David, 1995. Mortality Decline, the Demographic Transition, and Economic Growth. Unpublished, University of Chicago, February.

Ramsey, Frank, 1928. A Mathematical Theory of Saving. *Economic Journal* **38**, 543-559.

Rebelo, Sergio, 1991. Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy* **99, 3**, 500-521.

Romer, Paul M., 1986. Increasing Returns and Long-Run Growth. *Journal of Political Economy* **94, 5**, 1002-1037.

Romer, Paul M., 1987. Growth Based on Increasing Returns Due to Specialization. *American Economic Review* **77, 2**, 56-62.

Romer, Paul M., 1990. Endogenous Technological Change. *Journal of Political Economy* **98, 5**, part II, S71-S102.

Rosen, Sherwin, 1988. The Value of Changes in Life Expectancy. *Journal of Risk and Uncertainty* **1, 3**, 285-304.

Schumpeter, Joseph A., 1934. *The Theory of Economic Development*. Cambridge MA, Harvard University Press.

Sheshinski, Eytan, 1967. Optimal Accumulation with Learning by Doing, in Karl Shell, ed., *Essays on the Theory of Optimal Economic Growth*. Cambridge MA, MIT Press, 31-52.

Solow, Robert M., 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* **70, 1**, 65-94.

Swan, Trevor W., 1965. Economic Growth. and Capital Accumulation. *Economic Record* **32**, 334-361.

Uzawa, Hirofumi, 1965. Optimal Technical Change in an Aggregative Model of Economic Growth. *International Economic Review* **6**, 18-31.

World Bank, 1993. *World Development Report 1999*. Oxford University Press, Oxford.