Does education at all levels cause growth?  
India, a case study

Sharmistha Self a, Richard Grabowski b,∗

a Department of Economics, College of Saint Benedict, St. John’s University, St. Joseph, MN 56374, USA  
b Department of Economics, Southern Illinois University, Carbondale, IL 62901, USA

Received 9 January 2003; accepted 17 March 2003

Abstract

This paper seeks to examine the impact of education on income growth in India for the time period 1966–1996. Education is broken down into the categories of primary, secondary, and tertiary. Time series techniques are used to determine whether education, for each category, has a causal impact on growth. Furthermore, the education variables are also broken down by gender and analysis is carried out to determine whether the causal results vary by gender. The results indicate that primary education has a strong causal impact on growth, with more limited evidence of such an impact for secondary education. Finally, the evidence is quite compelling that it is female education at all levels, that has potential for generating economic growth.

© 2003 Elsevier Ltd. All rights reserved.

JEL classification: I21; 053

Keywords: Economic development; Human capital; Economic impact

1. Introduction

Traditionally, economic theory has emphasized physical capital accumulation as the most robust source of economic growth, at least in the short-run, with exogenous technical change being the long-run determinant of growth. However, attempts to make the long-run source of growth endogenous, rather than exogenous, led to the emergence of the concept of endogenous growth (Lucas, 1988). This literature has emphasized the importance of human capital as an endogenous factor of production to explain economic growth. Existing growth literature accepts education as one of the primary components of human capital since education, other than improving productivity of labor, has certain spillover benefits meaning that over and above benefiting the individuals who receive it, it also benefits society.

There is an abundance of empirical literature linking human capital to growth. The general consensus points at a positive relationship between growth and schooling though some recent papers have questioned this link leading to some research into the reliability of some of the available aggregate evidence (Temple, 2001). In comparison, there is even more limited and somewhat unclear evidence concerning the significance and relevance of different education levels to the growth mechanism. Moreover, much of the research related to human
capital is mostly based on evidence provided by cross-country studies, which have not concurred on the measure of education. This study considers the education–growth relation over a period of time in one country, with a focus on different levels of education (primary, secondary, and tertiary) and utilizing different measures of education. In addition, this study also analyzes the impact of education on growth by gender. Though some cross-country empirical studies imply a strong link between female education and economic development, a consensus has yet to be found. The focus of this paper, as the title suggests, is the impact of different education levels on India’s economic growth. Over and above studying the correlations between the variables of interest, the following specific premises are tested: (1) whether changes in education are responsible for or cause changes in economic growth (presumed in existing empirical work but, to our knowledge, not formally tested); and (2) do the relationships tested above change when we segregate the population by gender.

The rest of the paper is organized as follows. In Section 2, the variables being used to measure education are discussed, Section 3 provides a brief background of the Indian economy and its educational status, Section 4 explains the methodology, Section 5 reports and discusses the results, and Section 6 concludes.

2. Variables and data

Various studies on human capital have used different variables as proxies for education. The main reason for this, other than the difference in conceptualizing education, has been the lack of reliable data, especially for developing countries. Typically, however, two different variables have been most commonly used to measure education: enrollment rates and ratios and average years of education, commonly referred to as human capital stock. In this paper, both these measures are incorporated as a proxy for a flow of human capital. Primary, secondary, and tertiary enrollment rates consist of the number of individuals enrolled at each level, regardless of their ages, as a percent of the total population of appropriate age people at each level. These are based on UNESCO’s classification of age group appropriate with education level. Next, the growth rate of human capital stock, measured as the change in the mean years of education at each level, is analyzed. The same is repeated for individual genders, though gender-specific data for enrollment rates at the tertiary level are unavailable.

The time period for the study covers 30 years, 1966–1996. Data are annual and made available from the World Development Indicators database 1998 and 2001, provided by the World Bank, and include the enrollment variables and real per capita gross domestic product at market prices (1995 constant US$). The Penn World Tables 5.6 provides annual data for physical capital per worker (1985 international prices). Data on human capital stock, measured as average years of education at a particular level, are taken from Barro and Lee (2001). Data on per capita GDP and the capital–labor ratio are available for each year. However, data on enrollment rates are available every 5 years between 1965 and 1975 and thereafter annually. For 1966–1969 and 1971–1974, an exponential growth rate is calculated between the first and fifth year and the interim years are interpolated assuming an exponential smoothing process. Data on human capital stock are provided quinquennially for the entire period and the remaining years are filled in using an exponential smoothing process. Measurement errors are therefore likely for the human capital stock measure but, apart from this being the only source of reliable data on the measure, it is included since it is a more popular a measure of education.

3. Some background

Gross enrollment data for primary, secondary, and tertiary schooling levels are utilized as one of the measures of education in this paper. Enrollment ratios are a useful measure of education, though they do have some limitations. For example, gross enrollment ratios are not limited by age requirements or repeaters, which has been criticized on the grounds of leading to overstatement. Indeed, enrollment ratios can exceed 100%. However, in a country where compulsory education is not enforced, net enrollment ratios based on specific ages lead to
greater measurement error by not including students who fall outside certain age guidelines.\footnote{Gajraj (1992) finds about 31 countries where at least 50% of the students in grade 1 are older than the prescribed age. Barro and Lee (2001) revise their “fill-in procedure” to include gross enrollment ratios rather than net ratios since they find less measurement errors associated with the gross enrollment ratios.} Fig. 1 depicts the primary, secondary, and tertiary gross enrollment rates for India. Enrollment rates above 100% indicate the enrollment of students outside the appropriate age level.

Enrollment rates segregated by gender are also included in this paper. Figs. 2 and 3 depict primary and secondary enrollment by gender, and bring out the marked differences between male and female enrollment rates in Indian schools. Unfortunately, gender based education data for tertiary enrollments are unavailable at this time. It is clear from the figures that there is a large and persistent difference between male and female enrollment ratios. This gap has diminished marginally over time, at the primary level, but has remained fairly constant at the secondary level.

Next, variables representing the stock of human capital are taken into consideration. Human capital stock is measured in terms of educational attainment, that is, the average years of a particular level of education of the population aged 15 and above. Fig. 4 shows that, on average, the number of years of education at each level is dismal. The largest years at any particular level and the fastest growing is the primary level followed by the secondary level, with tertiary averaging the fewest number of years. It can be seen from the above figure that the human capital stock measure is typically lower than the enrollment rates. One possible explanation for this
could be that new entrants to the labor force are only a small fraction of those in work, hence, even large changes in enrollment rates take a much longer time to affect the average attainment level of the average population to any noticeable level.

Human capital stock measures are also provided for each gender. Figs. 5–7 show these for primary, secondary, and tertiary level, respectively. These three figures show that regardless of level, the average years of education at any particular level is higher for males as compared to females. Additionally, the growth of human capital stock also appears to be higher for males except at the primary level. Thus, the basic difference between genders is obvious with both enrollment rate measures and human capital stock measures.

4. Methodology

The basic definition of causality we use here is that defined by Granger (1969). Granger defined causality as testing whether lagged information on a variable $X$ provides any statistically significant information about a variable $y$ in the presence of lagged $y$. Though Granger causality has some limitations in application, such as measurement errors (Newbold, 1978) or temporal aggregation (Granger & Newbold, 1986), it needs to be understood that what is important for making a correct interpretation is the existence of some convincing theory for applying the causal mechanism. The existence of such a relation between education and growth has long been established and studied and extending it to different levels of education, to study whether or not this relation is causal, seem logical.

In order to determine the causal relationship between education and economic growth, the following hypothesis is tested,

$$H_0: E(\Delta y_t | \Delta y_{t-1}, \Delta y_{t-2}, \Delta y_{t-3}, \ldots, \Delta y_{t-n}, X_t, X_{t-1}, X_{t-2}, \ldots, X_{t-m}, Z_t, Z_{t-1}, Z_{t-2}, \ldots, Z_{t-m}) = E(\Delta y_t | \Delta y_{t-1}, \Delta y_{t-2}, \Delta y_{t-3}, \ldots, \Delta y_{t-n}),$$

for all $m > 0$—No Granger causality. Against the alternative of,

$$H_1: E(\Delta y_t | \Delta y_{t-1}, \Delta y_{t-2}, \Delta y_{t-3}, \ldots, \Delta y_{t-n}, X_t, X_{t-1}, X_{t-2}, \ldots, X_{t-m}, Z_t, Z_{t-1}, Z_{t-2}, \ldots, Z_{t-m})$$

for $m > 0$—Granger causality.
\[ \Delta Z_{t-1}, \Delta Z_{t-2}, \Delta Z_{t-3}, \ldots, \Delta Z_{t-m} \]
\[ \neq E(\Delta y_t | \Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-m}, \Delta Z_{t-1}, \ldots, \Delta Z_{t-m}) \]

For some \( m > 0 \)—Granger causality.

Here the \( \Delta y \) represents first difference of the log of per capita GDP, \( \Delta Z \) represents first difference of the log of the capital labor ratio, and \( \Delta X \) represents the first difference of the log of the education variables for each education level. It may be a concern that each education level is individually analyzed in each equation while allowing the constant term to account for all other influences. This, while introducing some bias in the results, increases the degrees of freedom while maintaining reliability of the results by limiting the number of explanatory variables (given the short time span being covered). Doing this results in some loss of generality, but it allows for the isolation and study of the causal impact of each individual education level and, at the same time, improves the degrees of freedom by reducing the number of regressors. When considering the enrollment rates, these are not further differenced to arrive at growth rates since they are in percent form. All other variables are represented in the first difference of their logs. The variables representing growth of the capital–labor ratio and per capita GDP, respectively, are common in each equation, keeping in line with a basic neoclassical production function relationship with human capital as an additional input.

In the gender-based analysis, the education variable for each level pertains to one gender at a time for reasons discussed above. This simplification avoids possible error from correlation between the gender-specific variables themselves and, as mentioned earlier, allows for the isolation and study of the impact of a specific gender at a specific level on growth. In this case, the null hypothesis above is modified by replacing the education variable with its gender-specific counterpart.

Before conducting any of the above tests, all of the relevant series are tested for stationarity, since standard inference procedures do not apply to regressions which contain an integrated dependent variable or integrated regressors. A formal method to test for stationarity of a series is the Unit Root test. To this effect the standard Augmented Dickey Fuller (ADF) test and the Phillips–Peron (PP) tests were utilized and all variables were found to be stationary. Next, the following model is formulated to test for a causal relation,

\[ \Delta y_t = \delta_0 + \sum_{j=1}^{m_1} \delta_j \Delta y_{t-j} + \sum_{j=1}^{m_2} \delta_j \Delta Z_{t-j} \]

\[ + \sum_{j=1}^{m_3} \delta_j \Delta X_{t-j} + e_{1t} \]

For the lagged variables appearing on the right-hand-side, the number of lags is determined using the Akaike Information Criterion (AIC) and Schwartz Criterion (SC) and the lag that gives the lowest AIC and SC and best fit is chosen. Adding lagged values of the dependent variable on the right-hand-side, other than fulfilling the Granger causality requirement, also reduces or eliminates the problem of spurious results due to serial correlation. A major part of the analysis depends on the choice of lag length since the results of the causality tests rely heavily on the time lags being imposed. If \( \delta_j \) and/or \( \delta_{ij} \) are found to be statistically significant and different from zero, we reject H0 and accept H1. In testing for the causal impact of gender based education on growth the above equation is modified as

\[ \Delta y_t = \delta_0 + \sum_{j=1}^{m_1} \delta_j \Delta y_{t-j} + \sum_{j=1}^{m_2} \delta_j \Delta Z_{t-j} \]

\[ + \sum_{j=1}^{m_3} \delta_j \Delta X_{t-j} + \sum_{j=1}^{m_4} \delta_j \Delta R_{t-j} + e_{1t} \]

\[ \Delta y_t = \delta_0 + \sum_{j=1}^{m_1} \delta_j \Delta y_{t-j} + \sum_{j=1}^{m_2} \delta_j \Delta Z_{t-j} \]

\[ + \sum_{j=1}^{m_3} \delta_j \Delta X_{t-j} + \sum_{j=1}^{m_4} \delta_j \Delta R_{t-j} + e_{1t} \]

where \( X_n \) is female education and \( X_m \) is male education. Eq. (2) represents the impact of female education at a particular level on growth and Eq. (3) represents the same for males. The only difference between the above equations and Eq. (1) is seen in Eq. (2) where an additional variable, \( R_t \), is added. \( R_t \) is the total fertility rate and measures the number of children that would be born to a woman if she were to live to the end of her child-bearing years and bear children in accordance with prevailing age-specific fertility rates. The total fertility rate is introduced as a distinguishing factor for females in order to analyze how the addition of this variable,

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]
along with education, affects the outcome of the study. Much work has been conducted to study, from a theoretical point of view, the role of gender and fertility on income or economic growth (Azariadis & Drazen, 1990, Becker & Barro, 1988, Blau & Kahn, 1992, Galor & Weil, 1996). These theories, based on either growth economics or family economics, have concluded that the gender-gap in education and the fertility rate have a negative impact on output per worker.9

In this paper, the analysis is also carried out with and without the inclusion of the total fertility rate in Eq. (2) in order to maintain uniformity and the results (presented in the following section) show that this has no significant effect. However, adding the total fertility rate as an additional explanatory variable seems to bring the analysis closer to reality since females attending school at all levels could be affected by child-bearing particularly in a country where, typically, students are seen to be over the age criteria and marriages usually occur at very low ages. Data on fertility are annual and provided by the World Bank’s World Development Indicator (1998 and 2001), though there are a few cases where the data are provided for alternate years and a weighted averaging is used to fill-in the missing years.

5. Results

Before carrying out the causality tests, correlations between the education variables and growth were estimated. Correlations provide an intuition regarding the relation between the variables and the direction of relation, which is not apparent from the causality regressions which are based on joint hypothesis tests. Here, both simple correlations and partial correlations were carried out.10 The results indicate significant positive correlations between the various levels of education and growth, whether one uses enrollment or human capital stock as the measure of education. The rest of this results section will be devoted to discussing the causality results.11

---

5.1. Primary education and growth

As explained earlier, two different education variables are tested for each level of education. Additionally, the same tests are conducted for each gender (for females with and without the inclusion of the total fertility rate). Table 1 provides results for primary education level. For simplicity, the results pertaining to the education variables alone are presented since this relation is the focus of interest.

From Table 1, it can be seen that primary education is not just strongly correlated with growth, but it has a strong causal impact on growth as well. This causal impact is felt by growth from both the primary enrollment rate variable as well as the change in average years of education at primary level. When the data are segregated by gender, the results do not change much, at least for females at primary level. For the female population at primary level, both variables, with or without the inclusion of the fertility rate variable, reflect a causal impact on growth. However, for the male population, this impact is seen only for enrollment rates, but not for the average years of primary education or human capital stock variable.

5.2. Secondary education and growth

Here, the method of analysis applied above is directed towards the study of secondary education. The results for the causal impact of secondary education on growth are presented in Table 2. These results show that there is some difference between the primary and the secondary level in terms of their impact on economic growth.

According to this table, while enrollment rates at secondary level continue to show a causal impact on growth, it is not the case with the human capital stock variable. The lack of any impact from human capital stock at the secondary level reduces the reliability of the estimate of the impact of the enrollment rate variable. Moreover, when the data are segregated by gender, it appears that the enrollment rate variable shows a causal impact on growth for males at a reduced significance level while the change in human capital stock of males at this level shows no causal impact. However, both the female enrollment rate variable (with and without the total fertility rate) and the change in human capital stock of females show a causal impact on growth (with and without total fertility rate). Thus, there is no ambiguity run connotation in the results. This method, however, applies to nonstationary data streams (usually data in levels) which are cointegrated. Since this paper looks at growth rates, VEC was not an option. Moreover, this methodology is also based on a joint hypothesis test and would not allow interpretation of magnitudes of individual variables.
Table 1
Causal relations at primary level

<table>
<thead>
<tr>
<th>Direction of causality: education to growth</th>
<th>Causality (F-statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>1.1 Enrollment</td>
<td>Yes (7.95)</td>
</tr>
<tr>
<td>1.2 Change in human capital stock</td>
<td>Yes* (3.24)</td>
</tr>
<tr>
<td><strong>Gender based</strong></td>
<td></td>
</tr>
<tr>
<td>1.3 Enrollment (male)</td>
<td>Yes (5.76)</td>
</tr>
<tr>
<td>1.4 Enrollment (female) w/o fertility</td>
<td>Yes (14.59)</td>
</tr>
<tr>
<td>1.5 Enrollment (female) w/ fertility</td>
<td>Yes (20.79)</td>
</tr>
<tr>
<td>1.6 Change in human capital stock (male)</td>
<td>No (0.006)</td>
</tr>
<tr>
<td>1.7 Change in human capital stock (female) w/o fertility</td>
<td>Yes (8.73)</td>
</tr>
<tr>
<td>1.8 Change in human capital stock (female) w/ fertility</td>
<td>Yes (8.93)</td>
</tr>
</tbody>
</table>

*Note: Yes* indicates 10% significance level while Yes indicates 5% significance level; each equation number refers to individual equations that were run including other variables such as an intercept, lagged values of the dependent variable, lagged values of the growth of capital labor ratio—detailed values are presented in an appendix available upon request from authors; all time lags are determined using the Akaike Information Criterion and the Schwartz Criterion.

Table 2
Causal relations at secondary level

<table>
<thead>
<tr>
<th>Direction of causality: education to growth</th>
<th>Causality (F-statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>2.1 Enrollment</td>
<td>Yes (6.13)</td>
</tr>
<tr>
<td>2.2 Change in human capital stock</td>
<td>No (1.57)</td>
</tr>
<tr>
<td><strong>Gender based</strong></td>
<td></td>
</tr>
<tr>
<td>2.3 Enrollment (male)</td>
<td>Yes* (3.03)</td>
</tr>
<tr>
<td>2.4 Enrollment (female) w/o fertility</td>
<td>Yes (6.53)</td>
</tr>
<tr>
<td>2.5 Enrollment (female) w/ fertility</td>
<td>Yes (7.31)</td>
</tr>
<tr>
<td>2.6 Change in human capital stock (male)</td>
<td>No (1.75)</td>
</tr>
<tr>
<td>2.7 Change in human capital stock (female) w/o fertility</td>
<td>Yes (3.44)</td>
</tr>
<tr>
<td>2.8 Change in human capital stock (female) w/ fertility</td>
<td>Yes (3.61)</td>
</tr>
</tbody>
</table>

*Note: Yes* indicates 10% significance level while Yes indicates 5% significance level; each equation number refers to individual equations that were run including other variables such as an intercept, lagged values of the dependent variable, lagged values of the growth of capital labor ratio—detailed values are presented in an appendix available upon request from authors; all time lags are determined using the Akaike Information Criterion and the Schwartz Criterion.

concerning the causal impact of female secondary education on growth.

5.3. Tertiary education and growth

The focus now shifts to tertiary level education and its relation with economic growth. Again, similar tests as the above are conducted. However, in this section, data on enrollment rates by gender are not available and, hence, the results on gender are dependent on the human capital stock variables alone.

According to Table 3, which presents results pertaining to the impact of tertiary education on growth, one does not see any causal impact in the general category. However, similar to the primary and secondary level, one finds evidence of a causal impact of the female population receiving tertiary education, but no such evidence exists for males. Thus, it appears that females, who are underrepresented in enrollment rates and in the accumulation of human capital stock at all education levels in India, are the ones having not just a strong correlation with the country’s growth, but having some predictive powers over growth as well (at all levels). For the population in general, however, the evidence points mainly at a causal relation for primary education with some weak evidence for secondary and none at all for tertiary education. With regard to the latter, however, caution must be exercised. Since the proportion of people undertaking such education is so low, it may be extremely difficult to pick up any causal effect from tertiary evaluation.
In this paper, the relationship between education at primary, secondary, and tertiary level and economic growth in India has been analyzed. According to the existing literature, there is a large amount of evidence for human capital having a significant impact on economic growth. In the present study, the same type of relation is seen in India in terms of correlations between education, at each level, and growth. However, correlations in themselves provide, at best, an intuition about the relation between two or more variables. Having found these encouraging correlations, this study utilized ‘Granger causality’ to analyze the predictive powers of each level of education on future growth in the presence of its own lagged values. Over and above allowing for a test of causality, this technique is helpful in time series regression analysis since it also helps to eliminate any possible serial correlation by adding lagged values of the dependent variable on the right hand side. The results showed that education, which in the correlation analysis indicated a strong positive relation between all education levels and growth, is causal only at the primary and secondary level.

This paper studied all these relations in terms of gender, that is, the casual impact of each gender at each education level on growth. The results showed that female education at all levels has potential for generating economic growth. Males, on the other hand, appear to have a causal impact on growth only at primary level and perhaps, weakly, at the secondary level.

In closing, a word of caution is in order. The conclusion that primary education is the main causal force in economic growth in India must be qualified since education’s impact is likely to show only after long time lags and there may be important omitted variables. Thus, further research using more extensive data sets is certainly required.

### References


