NBER WORKING PAPER SERIES

INTERNATIONAL COMPARISONS OF EDUCATIONAL ATTAINMENT

Robert J. Barro

Jong-Wha Lee

Working Paper No. 4349

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1993

Harvard University and International Monetary Fund, respectively. We are grateful to Birger Fredriksen and George Psacharopoulos for helpful discussions and to the Division of Statistics on Education, UNESCO, for the provision of information. Any opinions expressed in this paper are those of the authors and not of the institutions with which they are affiliated. This paper is part of NBER's research program in Growth. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #4349
April 1993

INTERNATIONAL COMPARISONS OF EDUCATIONAL ATTAINMENT

ABSTRACT

Many theories of economic growth stress the role of human capital in the form of education, but empirical studies have been hampered by inadequate data. We describe a data set on educational attainment that we have constructed for 129 countries over five-year periods from 1960-1985. We use census/survey information to fill over 40% of the cells, and we use school-enrollment figures in a perpetual-inventory framework to fill the remainder. The data refer to male and female attainment of the adult population at four levels: no-schooling, primary, secondary, and higher. We also provide a rough breakdown into incomplete and complete attainment at the three levels of schooling. We then take account of cross-country variations in the durations of schooling at each level to provide figures on total years of attainment.

Robert J. Barro
Department of Economics
Littauer Center 120
Harvard University
Cambridge, MA 02138
and NBER

Jong-Wha Lee
Korea University
Department of Economics
Anam-Long, Sunbuck-ku
Seoul 136-701, Korea

Many theoretical models of economic growth, such as those of Lucas (1988); Becker, Murphy, and Tamura (1990); Rebelo (1992); and Mulligan and Sala–i–Martin (1992), emphasize the role of human capital in the form of educational attainment. Empirical studies of growth for a broad cross section of countries, such as those by Romer (1990), Barro (1991), and Kyriacou (1991), have used proxies for human capital. These studies have, however, been hampered by the limited educational data that are available on a consistent basis for these kinds of cross sections.

We describe in this paper the data set on educational attainment that we have constructed for 129 countries over five-year periods from 1960 to 1985. Our concept of human capital is the years of completed schooling for persons aged 25 and over. The underlying information comes from census and survey estimates of educational attainment from UNESCO Statistical Yearbooks, Kaneko (1986), U.N. Demographic Yearbooks, and some other sources. These data appear to be reasonably consistent over time and across countries.

The available data motivated us to focus on educational attainment for the population aged 25 and above, rather than for a younger age category or for subgroups of the population such as the labor force or employed persons. Within the 25 and over category, we construct figures for years of education at seven levels: no schooling, incomplete primary schooling, complete primary schooling, first cycle of secondary schooling, second cycle of secondary schooling, incomplete higher education, and complete higher education. The data are, however, more accurate for the four major levels of attainment: no schooling, some primary, some secondary, and some higher education.

The basic data set has a maximum of 774 observations on each variable—6 time-series values for 129 countries. Roughly 40% of these observations for the four major levels of attainment are drawn from figures presented in censuses or surveys (most of

which are full censuses). In addition, we used information on adult illiteracy to proxy for the no-schooling category for some cases in which census/survey figures were unavailable. We fill in most of the remaining cells for the four-level classification from an estimation method that exploits the available data on school enrollment rates and population by age. The breakdown into the seven levels of schooling is based on limited information on completion ratios for primary, secondary, and higher education.

We have also used the same sources and methods to construct estimates of female educational attainment. This information allows us to compare the attainment of females with that of males across countries and over time.

Our data measure years of school attainment, but do not adjust for quality of education, length of school day or year, and so on. The necessary information to make these kinds of adjustments do not seem to be available for the broad cross section of countries that we are considering, although it would be possible to take account of elements such as public expenditures on education and pupil-teacher ratios.

Although the lack of adjustment for quality of education is an important omission, we believe that our data set nevertheless provides information on the stock of human capital that is superior to that available before. Some previous empirical studies have, for example, used school enrollment ratios and adult literacy rates. These variables are available for a large number of countries, but do not measure accurately the stock of human capital that is available for current production. School enrollment ratios reflect current flows of education, and the accumulation of these flows will be one element in the stocks of human capital that will be available later. Adult literacy rates measure one component of the current stock of human capital, but do not reflect the skills that are obtained beyond the most elementary levels of schooling.

The inadequacy of school enrollment ratios and adult literacy rates has motivated other researchers to construct more appropriate measures of the stock of human capital.

Psacharopoulos and Ariagada (1986a, 1986b, 1991) have used census and survey data to compile information about the educational attainment of the labor force (or, in some cases, of the adult population). The main shortcoming of their data set is that its coverage is too small: most countries have only one time-series observation, and the year covered differs across the countries.

Kyriacou (1991) has constructed panel estimates of educational attainment for a large number of countries. He relates the available census figures from Psacharopoulos and Ariagada for years in the 1970s to school enrollment ratios. He then extrapolates this relationship to other years by using the data on school enrollment ratios. Our procedure should be more accurate because we begin by assembling a much larger number of census/survey data points and then use school enrollment ratios to fill in on a much more limited basis.

Lau, Jamison, and Louat (1991) also provide panel estimates of educational attainment. They cumulate flows of schooling based on the school enrollment data and on assumptions about survival rates of the population. The accuracy of their estimates is an issue because they do not use census benchmarks for starting or intermediate values of educational stocks.

Section I discusses the data on school enrollment ratios, adult literacy rates, and census/survey estimates of educational attainment. We then summarize the estimation procedures employed by Psacharopoulos-Ariagada, Kyriacou, and Lau-Jamison-Louat. Section II describes our method for estimating educational attainment of the population aged 25 and above. We discuss the basic data, and we detail our procedure for filling in the missing observations. Section III highlights the main features of the complete data set, and Section IV provides a comparison with alternative measures of human capital. Section V discusses our estimates of female educational attainment, and Section VI provides some concluding observations.

I. Statistical Measures of Human Capital

School Enrollment Ratios

School enrollment ratios are widely available across countries and have been used in numerous studies. (See Tilak (1989) for a survey of these studies and see Fredriksen (1991) for a discussion of the underlying data.) Enrollment ratios are available for three levels of schooling: primary, secondary, and higher. Gross enrollment ratios relate the total number of students at a given level to the population of the age group that national regulation or custom dictates would be enrolled at that level. Net enrollment ratios modify the numerator of the ratio to count only the students enrolled within the designated age group. For example, for a country in which children start primary school at age six and in which primary school has six grades, the gross and net enrollment ratios for primary education are given by

$$\text{gross enrollment ratio} = \frac{\text{total enrollment in primary school}}{\text{total population aged 6—11 years}}$$

$$\text{net enrollment ratio} = \frac{\text{total enrollment aged 6–11 in primary school}}{\text{total population aged 6–11 years}}$$

Thus, the net enrollment ratio is between zero and one, whereas the gross enrollment ratio can exceed one. Although the net enrollment ratio is more appropriate for gauging the accumulation of human capital, the gross ratio has usually been used because it is more often available for developing countries.

School enrollment ratios have several deficiencies as measures of stocks of human capital. First, the current enrollment ratios measure the flows of schooling; the cumulation of these flows creates the future stocks of human capital. Because the

educational process takes many years, the lag between flows and stocks is long (see Psacharopoulos and Ariagada (1986b, pp. 1–2) for a discussion). If the appropriate lag is considered, then the construction of human capital stocks still requires an estimate of initial stocks. Errors are also introduced because of mortality and migration and because the net enrollment ratios are often unavailable. The gross enrollment ratios introduce errors related to repetition of grades and dropouts, phenomena that are typically high in developing countries.[1]

Another problem is that the underlying data on school enrollment are of doubtful quality for developing countries. Most information collected by UNESCO comes from annual surveys of educational institutions in each country. The typical practice is that the person responsible for administering each institution answers a number of questions about his or her institution. Chapman and Boothroyd (1988) note "... in several countries, headmasters have been observed to inflate reported enrollment based on their experience that higher enrollment figures lead to more resources (supplies, textbooks, budget) being allocated to the school." In Yemen Arab Republic, for example, the enrollment reported by school officials to the Ministry of Education in 1985 exceeded by 56% the numbers recorded in the school files. Thus, in general, the reported enrollment rates have an upward bias.

An additional source of upward bias in the enrollment figures is that the data refer to the registered number of students at the beginning of each school year. The actual number of children that attend school during the year can be substantially lower. The error is particularly serious for developing countries in which the government punishes parents that do not register their children at primary schools. In rural areas of China,

---

[1]For the total of developing countries in 1980, Fredriksen (1983) estimates that the average gross enrollment ratio at the primary level was 86%. The elimination of repeaters reduces the estimated value to 73%.

for example, Fredriksen (1991) estimates that actual school attendance was 30% below the registered number for the first grade.

Adult Literacy Rates

The adult literacy rate is also widely available and has frequently been used in empirical studies. For example, the United Nations uses these data to construct an index of human capital (United Nations Development Programme [1990]), and Romer (1990) used them to estimate the relation between human capital and economic growth.

One desirable feature of the adult literacy rate is that it measures a stock of human capital for the adult population, whereas the school enrollment ratios measure flows of education. The literacy rate is, however, less widely available because the underlying information typically comes from general population censuses and surveys, activities that usually occur only once per decade.

The concept of adult literacy has some problems with respect to international comparability. Since 1958, UNESCO has provided a standardized definition, which can be summarized as "... a person is literate who can both read and write a short simple statement of his everyday life" (Carceles [1990, p. 6]). This definition is not easy to apply systematically to determine whether a person is literate. Since the evaluation is not based on any objective tests, the measures of literacy could be biased in a way that varies across countries and over time.

A more basic problem is that literacy is only the first stage in the path of human capital formation. Other aspects of human capital that are important for labor productivity include numeracy, logical and analytical reasoning, and various types of technical knowledge. If literacy is used to measure the stock of human capital, then the implicit assumption is that education beyond the most elementary level does not contribute significantly to productivity.

We have found in practice that the reported figures on adult illiteracy are closely correlated with census/survey data on the fraction of the adult population who have no school attainment. For 158 observations, the correlation between the no-schooling value reported by a census or survey and the adult illiteracy rate is 0.95. Given this high correlation, we decided to use adult illiteracy rates to proxy for the no-school percentage for some cases in which census/survey figures were unavailable.

Educational Attainment

Census/survey data provide information on levels of educational attainment for a specified population group, such as the labor force or persons aged 25 and over. Psacharopoulos and Ariagada (1986a, 1986b) have used national census publications to compile data on educational composition of the labor force for 99 countries. The levels of education cover five categories above no school: incomplete primary, complete primary, incomplete secondary, complete secondary, and higher. Psacharopoulos and Ariagada estimate the average years of schooling from the formula

$$\text{average years of schooling} = \sum_j YR_j \cdot HS_j$$

where j is the schooling level (incomplete primary, ...), YR_j is the number of years of schooling represented by level j, and HS_j is the fraction of the population for which the jth level is the highest value attained.

Although this concept seems superior to enrollment ratios or literacy rates as a measure of human capital, some shortcomings nevertheless apply. The definition of the labor force varies across countries, and Psacharopoulos and Ariagada are in any case forced by the available data to use the population of male adults instead of the labor

force for many countries (see appendix B in Psacharopoulos and Ariagada [1986b]). Also, for many purposes—such as studies of the impact of female educational attainment on fertility, health, and school enrollment—it would be undesirable to ignore the education of persons who are outside the formal labor force.

The key drawback, however, of the Psacharopoulos-Ariagada data set is the small sample size. Information is provided only for isolated points in time for each country, and only 34 countries have more than one observation.

The difficulty in defining and measuring the labor force suggests that educational attainment for other subsets of the population may be more accurate and therefore more useful. Kaneko (1986) compiled data on the educational composition of the adult population by age and sex for 78 countries since 1970. UNESCO (1983) provided a more comprehensive set of the same statistics for 149 countries and territories. It is surprising that these data have not yet been used much in empirical research on economic growth.

The limited number of observations for census/survey measures of educational attainment has motivated some researchers to construct alternative estimates of human capital. Lau, Jamison, and Louat (1991) used the annual data on school enrollment ratios (already discussed) to estimate time series of total stocks of educational attainment for the working-age population. They began by extrapolating their data on gross enrollment ratios, applying from 1950 to 1980, to the periods 1900–50 and 1980–85. They then used a perpetual inventory method to construct a time series for educational attainment at the primary and secondary levels. The estimation assumed zero starting stocks in 1900 (a year far enough back for the starting point not to matter) and used age-specific survival rates based on regional data. One serious drawback of this procedure is that it does not exploit the available statistics on benchmark stocks of educational attainment. This lack of benchmarking probably explains why the resulting estimates of educational attainment turn out not to be highly correlated with the census

figures provided by Psacharopoulos and Ariagada (1986a)—see the appendix to Lau, Jamison, and Louat (1991).

Kyriacou (1991) has estimated educational attainment by combining the census information with the data on school enrollment. He regressed the census data from Psacharopoulos and Ariagada (1986a) for 42 countries in the mid 1970s on prior values of school enrollment ratios. He then used this relationship to estimate educational attainment for other years and other countries. Thus, he assumed that the relationship was stable over time and across countries. The estimates are, not surprisingly, closely correlated with the original census data in the mid 1970s. The relationships turn out, however, to be much less accurate for other time periods. This kind of problem was discussed earlier in UNESCO (1978). They noted that projections of educational attainment based on the relation between attainment and enrollment ratios were not reliable because the relation was unstable over time.

II. A New Data Set on Educational Attainment

We now describe the procedures for assembling our data base. We begin with the census/survey data on educational attainment and then discuss our method for estimating the missing observations.

The Basic Data

Our basic data on educational attainment come primarily from issues of UNESCO *Statistical Yearbook*, which reports census and survey data by age and sex. Most of these data come from full censuses. We focus here on information for the educational attainment of the total population aged 25 and over. The data on the educational attainment of the female population aged 25 and over are discussed in section V.

About two-thirds of the census/survey data reported by UNESCO are U.N. figures, and the remainder come from national sources (see UNESCO [1983]). Additional observations in our data set come from U.N. *Demographic Yearbook*, 1983; Kaneko (1986); various issues of *Statistiches Jahrbuch* (for Germany), and Taiwanese government publications (for Taiwan). This set of basic data exhausts the census/survey figures on educational attainment that we know of.

We focus on the population aged 25 and over in order to obtain the widest possible coverage. Thus, the data differ from those presented by Psacharopoulos and Ariagada, who considered the educational attainment of the labor force.[2] For the population aged 25 and over, we have obtained at least one observation since 1945 for 129 countries, including 10 formerly centrally-planned economies. (Two of the more prominent missing countries are China and Nigeria.) Most countries, especially those outside of sub-Saharan Africa, have three or more observations. Thus, the data set provides a reasonable basis for panel estimation.

Tables 1 and 2 summarize the coverage of the basic data. Table 1 shows that 77 of the 129 countries have 3 or more observations. Table 2 shows the number of countries with data for the indicated survey year from 1945 to 1985. (If the original survey date does not coincide with one of these points, then we rearranged the observation to apply to the nearest 5-year value; for example, any date between 1968 and 1972 is classed as 1970.) The table shows that over 70 observations are available for the census years 1960, 1970, and 1980. The total number of observations is 364, of which 311 are between 1960 and 1985. These 311 observations fill 40% of the 774 possible cells for 129 countries from 1960 to 1985.

---

[2]We already mentioned that Psacharopoulos and Ariagada are forced to use different definitions of the labor force in each country. For 33 countries, they use the male population aged 15 and over as a proxy for the labor force. In these cases, their data come from the same sources as the UNESCO data that we use.

We have obtained data that correspond to the six levels of educational attainment established by UNESCO's "international standard classification of education," as follows:

No schooling: Anyone who completed less than one year of primary school. For a few countries in which the data on the breakdown between no schooling and primary attainment were unavailable, we used the fraction of adults classed as illiterate to measure the no-schooling percentage. Appendix Table A.1 uses the symbol # to denote these cases. Cases in which the UNESCO report indicates that they used illiteracy data to compute the no-schooling percentage are denoted by the symbol &.

Incomplete first level: Anyone who received at least one year of primary education but did not complete the last year of primary school.

Complete first level: Anyone who completed the final year of primary school (or, in some cases, reached the penultimate year of primary school), but did not advance to secondary school.

Entered second level, first cycle: Anyone who entered the lower stage of secondary school but did not advance to the higher stage of secondary school.

Entered second level, second cycle: Anyone who entered the higher stage of secondary school but did not advance to post-secondary studies.

Higher level : Anyone who entered post-secondary school.

Appendix Table A.1 shows the complete set of census/survey data for each country.[3] For convenience of presentation, we classify the sum of incomplete and complete first-level education as *First Level Total* and the sum of first and second cycle second-level education as *Second Level Total*. Table A.1 shows that for the majority of countries the census/survey data do not distinguish incomplete from complete primary or the first cycle from the second cycle of secondary education.

As already mentioned, the no-schooling percentages are highly correlated with reported figures on adult illiteracy—the correlation for 158 observations is 0.95.[4] We have therefore used the data on adult illiteracy to estimate the no-schooling percentage for cases in which census/survey data on educational attainment are unavailable. We have used the illiteracy numbers only for countries that also report adult illiteracy and a census/survey value for the no-schooling percentage in another year. We then the use the ratio of no-schooling to illiteracy in this other year[5] to adjust the illiteracy number for the year in which the no-schooling fraction is unavailable. This procedure yields 124 additional observations on the no-school category, most of which are for Africa and Latin America. More than half (68) of the observations are for 1985. These values appear in Table A.2 under the no-schooling column and are marked by the symbol #.

We have used the available census/survey data as benchmarks and then estimated the missing observations for each country from 1960 to 1985. We carry out the estimation in two steps: first, we estimate the missing observations at the broad four-level classification: no schooling, first level total, second level total, and higher; and

---

[3]The data set includes 43 observations that we used for which the covered age group includes persons younger than 25 (see Appendix Table A–1). We have not yet adjusted our estimates to take account of this discrepancy, although our preliminary computations suggest that the necessary adjustment is minor.

[4]Some of the figures on the no-schooling fraction in the census/survey reports were themselves based on estimates of adult illiteracy. The 158 observations are those for which the no-schooling value was generated independently of the literacy number.

[5]If more than one other year is available, then we use the closest year.

second, we estimate the breakdown of first and second level education into their two sub-categories. We also provide a rough disaggregation of higher-level attainment into incomplete and complete schooling by using the limited information provided by Kaneko (1986).

Estimation of Missing Observations at the Broad Four-Level Classification

At each of the four broad levels, the available census/survey data fill 40% of the possible cells (311 out of 774) for the 129 countries from 1960 to 1985. We have filled an additional 124 cells for the no-schooling category by using the adult illiteracy numbers. We now describe our methodology for using data on school enrollment ratios and population by age to fill in the remaining cells.

We have the necessary data on gross school enrollment ratios and population by age from 1950 to 1985 for 116 of the 129 countries that have at least one census/survey observation on educational attainment.[8] We are missing the enrollment data for 13 countries—Antigua and Barbuda, Belize, Brunei, Dominica, Namibia, Seychelles, Solomon Islands, St. Kitts, St. Lucia, St. Vincent, Vanuatu, United Arab Emirates and Western Samoa. Therefore, our fill-in procedure applies to 116 countries from 1960 to 1985. For 10 of these countries—Benin, Cameroon, Central African Republic, Congo, Egypt, Gambia, Guinea–Bissau, Mali, Rwanda, and Yemen—we ultimately concluded that the fill-in procedure was unsatisfactory for at least part of the sample, probably

---

[8]The data on gross enrollment ratios are from various issues of UNESCO, *Statistical Yearbook*. We obtained these data at five-year intervals from 1950 to 1980. If the number for the five-year value was missing, then we used a value reported for a nearby year; for example, we would use a value for 1966 to represent 1965. If we proceed in this manner, then the data set on enrollment ratios is nearly complete for the 116 countries from 1950 to 1980. For most countries, we estimated primary school enrollment ratios for 1945 (needed to fill-in the attainment data for 1960) by extrapolating the observed values for 1950 and 1955. For a few countries, we had to interpolate to fill in other missing values; for example, an average of 1960 and 1970 would be used if necessary to represent 1965. If the reported value of a gross enrollment ratio exceeded 1.0, then we substituted 1.0.

because of gross errors in the school-enrollment data. Therefore, the full time series of six observations applies to only 106 countries.

Our main procedure is a perpetual inventory method that starts with the census/survey figures as benchmark stocks and then uses the school enrollment ratios to estimate changes from the benchmarks. Let L_t be the population aged 25 and over at time t and H_{jt} be the number of people within this adult population for whom j is the highest level of educational attainment. We let j=0 for no school, j=1 for total primary, j=2 for total secondary, and j=3 for higher. Let $h_{jt} \equiv H_{jt}/L_t$ be the proportion of the adult population for whom j is the highest level attained. We have survey estimates of h_{jt} for various years and countries, and we seek to estimate the missing values.

Let $PRI_{t-\tau}$ be the gross enrollment ratio for primary school, $SEC_{t-\tau}$ the ratio for secondary school, and $HIGH_{t-\tau}$ the ratio for higher education, all observed at time $t-\tau$. Let $L25_t$ be the population aged 25–29 at time t—this number represents the people who entered into L_t, the overall population aged 25 and above, during the last five years.[7] We assume that these new entrants into the adult population would have received primary education 15 years earlier (if they attended primary school), secondary education 10 years earlier (if they attended secondary school), and higher education 5 years earlier (if they received higher education).

We now illustrate our procedure in detail for the no-schooling category. The estimated number of persons aged 25 and over who have no educational attainment at time t is given by

$$(1) \qquad H_{0t} = H_{0,t-5} \cdot (1-\delta_t) + L25_t \cdot (1-PRI_{t-15})$$

[7]The data on population by age are from United Nations (1991).

where δ_t is the the proportion of people aged 25 and over in year t–5 who did not survive to year t.[8] We assume that the fraction of the population aged 25 to 29 who have no elementary education equals the fraction, $1-PRI_{t-15}$, who were not enrolled in primary school 15 years before. This assumption could be incorrect because of errors in the gross enrollment figures, because of children moving in and out of primary school, because of timing problems, because survival probabilities over the previous 15 years are correlated with educational attainment, and because of migration.

We estimate the death probability that appears in equation (1) from

$$(2) \qquad \delta_t \simeq (L25_t + L_{t-5} - L_t)/L_{t-5}$$

This formula neglects any mortality for those persons who were aged 20–24 five years previously and again ignores migration. If we substitute from equation (2) into equation (1) and rearrange terms, then the formula for the fraction of the population with no schooling becomes

$$(3) \quad h_{0t} \equiv H_{0t}/L_t = [1-(L25_t/L_t)] \cdot h_{0,t-5} + (L25_t/L_t) \cdot (1-PRI_{t-15})$$

Hence, the estimated value of h_{0t} is a weighted average of the prior value, $h_{0,t-5}$, and the fraction of the population, $(1-PRI_{t-15})$, who were not enrolled in primary school 15 years previously. The weight on the schooling experience of these new entrants is $L25_t/L_t$, the fraction of the population currently aged 25–29.

---

[8]We neglect here any mortality for persons aged 20-24 five years previously, and we assume that the survival probability for persons who were 25 and over is independent of the level of educational attainment. Some error is introduced here if educational attainment is growing rapidly because the older people then have less human capital and a greater probability of dying.

The procedures for the other levels of educational attainment are analogous. The resulting formulas are

$$(4) \quad h_{1t} \equiv H_{1t}/L_t = [1-(L25_t/L_t)]\cdot h_{1,t-5} + (L25_t/L_t)\cdot(PRI_{t-15}-SEC_{t-10})$$

$$(5) \quad h_{2t} \equiv H_{2t}/L_t = [1-(L25_t/L_t)]\cdot h_{2,t-5} + (L25_t/L_t)\cdot(SEC_{t-10}-HIGH_{t-5})$$

$$(6) \quad h_{3t} \equiv H_{3t}/L_t = [1-(L25_t/L_t)]\cdot h_{3,t-5} + (L25_t/L_t)\cdot HIGH_{t-5}$$

Note that equations (3)–(6) imply that the constructed values, h_{0t}, h_{1t}, h_{2t}, and h_{3t}, must add to 100% (assuming that the initial values, $h_{0,t-5}$, $h_{1,t-5}$, $h_{2,t-5}$, and $h_{3,t-5}$, added to 100%).

Accuracy Test

The perpetual-inventory method just described has several possible sources of error, one of which is inaccuracies in the underlying data on gross enrollment ratios. To evaluate the use of this method to estimate missing observations, we now assess its accuracy for the 30 countries that have complete census estimates for 1960, 1970, and 1980. We can use the benchmark values from 1960 and then use equations (3)–(6) in the forward direction to estimate attainment in 1970. (We adjust here to consider a 10–year interval rather than a 5–year interval.) Similarly, we can use the benchmark values from 1970 to estimate attainment in 1980. We call these values *forward-flow* estimates. We can also start with benchmark values in 1970 or 1980 and use the equations in the backward direction to estimate attainment in 1960 and 1970, respectively. We call these values *backward-flow* estimates.

We also compare the accuracy of the forward- and backward-flow estimates with forecasts derived from simple linear trends: extrapolations from the values for 1960 and 1970 to an estimate for 1980 and from the values for 1970 and 1980 to an estimate for 1960. We also construct linear interpolations from the values for 1960 and 1980 to estimates for 1970.

We carried out a simulation exercise in which we regressed the observed values of the various levels of educational attainment in 1960, 1970, and 1980 for the 30 countries on the estimates generated from forward- and backward-flow methods and from linear extrapolations and interpolations. For the 1960 observations, we included the backward-flow estimate from 1970 and the linear extrapolation from 1970 and 1980. For 1970, we used the forward-flow estimate from 1960, the backward-flow estimate from 1980, and the linear interpolation from 1960 and 1980. For 1980, we included the forward-flow estimate from 1970 and the linear extrapolation from 1960 and 1970.[9]

The results indicated that first, the linear extrapolations for 1960 and 1980 were insignificant in all cases; second, the backward-flow estimate for 1970 was insignificant in all cases; third, the forward-flow estimate was significant in all cases for 1980; fourth, the forward-flow estimate and the linear interpolation for 1970 were jointly significant in all cases; and fifth, the backward-flow estimate was significant for all cases in 1960, a situation in which the forward-flow estimate was not present.

We were guided by these results to fill in the missing observations as follows. We used a forward-flow estimate if this estimate were available and if it was infeasible to interpolate between two observations. These situations correspond to the results for 1980 from the simulations. If the forward-flow estimate was unavailable, then we used a

---

[9]We estimated subject to the restriction that the educational attainment variable, h_{jt}, was between zero and one for all levels j and that the sum of the h_{jt} over the four levels j equaled one.

backward-flow estimate only based on a later benchmark value for attainment. These situations correspond to the results for 1960 from the simulations. (Note that, if a country has at least one observation on attainment and a full set of values of enrollment ratios, then either the forward-flow or backward-flow estimate must be available.)

For cases in which it was possible to interpolate between two observations on attainment, the simulations for the 30 observations for 1970 revealed that the best fit was the weighted average, .40·(forward-flow estimate) + .60·(linear interpolation). (These coefficients came from joint estimation over all levels of educational attainment.) We used this formula when feasible to fill in the missing observations on attainment at all levels of education.[10]

Table 3 provides some standard quantitative measures of the accuracy of these procedures for the 30 countries that have data on attainment for 1960, 1970, and 1980. The table shows for the four levels of schooling and for each year the actual mean of attainment, the root-mean squared error from the estimation procedure, and Theil's U statistic.[11] The numerator of this statistic is the root-mean squared error. The denominator is the sum of the squares of the actual values plus the sum of the squares of the estimated values. The U Statistic is analogous to $1-R^2$ in that the best possible fit corresponds to U=0 and the worst possible fit to U=1. Although the estimated values

---

[10]The filled-in values add to 100% by construction (see equations [3]–[6]). However, for cases in which we use the illiteracy rate data to estimate the no-school percentage, the figures need not add to 100%. We adjusted only the primary school category in these cases so that the categories added to 100%. We adjusted in this manner because the discrepancy arises mainly for less-developed countries for which most of the population is in the no-school and primary categories and for which the enrollment figures at the primary level tend especially to be exaggerated. It is also possible that a filled-in value would be negative, an outcome that arose in practice mainly for backward-flow estimation of primary attainment for some less-developed countries. We have omitted the data for these years from Table A.2 in these cases. In a few other situations, the filled-in values at the secondary and higher levels were negative but very small in magnitude. We have set these values at 0.1 and included the figures in Table A.2.

[11]See Pindyck and Rubinfeld (1981, pp. 364–365) for a discussion.

explain a large fraction of the variation of the actual values—the U statistics range between 0.09 and 0.36—the table makes clear that the remaining measurement error is substantial.

Estimation of Sub–Categories of Educational Attainment

We now describe our procedure for estimating missing observations for the sub-categories of each education level—primary, secondary and higher. First, we describe the procedure for the primary schooling category.

We filled in the missing observations using information from the available census/survey data. The data presented in Table A.1 show that 165 observations from 94 countries are available for the breakdown of first-level attainment into incomplete and complete categories—see Table 4 for a summary of the available observations. We call the ratio of complete primary schooling (the fraction of the population over 25 who completed primary school but did not enter secondary school) to total primary schooling (the fraction who entered primary school but did not advance to secondary school) the completion ratio for primary education. We assume that this ratio is determined by the characteristics of each country and by features of the region to which the country belongs. Accordingly, we regressed the observed completion ratios on 5–year or 10–year lagged values and on regional dummies:

$$(7) \qquad R_{it} = a + b \cdot R_{i,t-5} + c \cdot (\text{region dummy}) + u_{it}$$

$$(8) \quad R_{it} = a(1+b) + b^2 \cdot R_{i,t-10} + c(1+b) \cdot (\text{region dummy}) + u_{it} + bu_{i,t-5}$$

where R_{it} is the completion ratio for primary education in country i in year t and (region dummy) denotes a set of regional dummies. Equation (8) was constructed by

substituting for $R_{i,t-5}$ from a lagged version of equation (7), $R_{i,t-5} = a + b \cdot R_{i,t-10} + c \cdot (\text{region dummy}) + u_{i,t-5}$. We proceed in this manner so that we can exploit observations on R_{it} for a single country that are spaced either at 5– or 10–year intervals.

Equations (7) and (8) were estimated jointly by weighted least squares. The number of observations is 17 for equation (7) and 43 for equation (8), that is, we have 17 and 43 cases in which a country has two observations that are spaced respectively at a 5–year and 10–year interval. We used the regression results to estimate missing observations in a forward direction. For example, for a given observed value of R_{it} in 1975, we can use the equations to estimate completion ratios for 1980 and 1985.

To fill–in in the backward direction, we used a regression of completion ratios on 5–year and 10–year lead values and on the regional dummies. For example, for a given observed value of R_{it} in 1970, we can use these equations to estimate completion ratios for 1965 and 1960.

We used these forward and backward methods to estimate primary completion ratios from 1960 to 1985 for the 94 countries that have at least one actual observation.[12] For the remaining countries, which lack data on the primary completion ratio, we used the regional means shown in Table 4.

We apply the same procedure to estimate completion ratios for secondary education, that is, the ratio of second-cycle attainment to total secondary attainment. Table A.1 shows that 116 observations from 75 countries are available for the breakdown of secondary attainment into first and second cycles. In this case, the joint estimation of equations (7) and (8) was carried out with 13 and 22 observations respectively. We used these regressions and comparable equations in the backward

---

[12]If forecasts were available from forward and backward estimation, then we used the linear combination of the two forecasts that was obtained by regressing the actual observations on the two forecasts.

direction to estimate secondary completion ratios for the 75 countries that had at least one actual observation. The regional means shown in Table 4 were used in the remaining countries.

We also made a rough estimate of the disaggregation of higher education into incomplete and complete components based on the limited information available in Kaneko (1986). He reports the UN censuses and surveys undertaken from 1975 to 1984 on educational attainment for two sub-categories of tertiary schooling. The lower tertiary level includes junior-college graduates and university attendees without degrees, and the upper tertiary level includes university graduates with degrees and post-graduates. These data comprise one observation for each of 37 countries, as shown in Table 5.

We used these numbers for the 37 countries to construct the completion ratio for higher education, that is, the ratio of upper-level attainment to total higher attainment. Then we assumed that the completion ratio remained constant from 1960 to 1985, that is, the lack of time-series data forced us to neglect any variations in the ratio over time. For the remaining countries, we used the average of completion ratios in the country's region.

III. The Complete Data Set

Appendix Table A.2 shows the full data set on educational attainment. The census/survey data from Table A.1 have been combined here with the filled–in observations that we described above. (An asterisk indicates that the figure comes directly from a census/survey observation.) For 106 countries, all cells from 1960 to 1985 have been filled in. As mentioned before, the data are incomplete for 23 countries. The table also includes observations for 1950 and 1955 if a census or survey value is available for that year.

Table A.2 also reports the average number of years of schooling at all levels attained by the population aged 25 and over. We construct the average years of schooling from the formula,

$$\begin{aligned}\text{average years of schooling} &= DUR_p \cdot [(1/2)h_{ip} + h_{cp}] + (DUR_p + DUR_{s1}) \cdot h_{is} \\ &+ (DUR_p + DUR_{s1} + DUR_{s2}) \cdot h_{cs} + [DUR_p + DUR_{s1} + DUR_{s2} + (1/2)DUR_h] \cdot h_{ih} \\ &+ (DUR_p + DUR_{s1} + DUR_{s2} + DUR_h) \cdot h_{ch}\end{aligned} \tag{9}$$

where each percentage refers to the fraction of the population for which the jth level of schooling is the highest attained: $j = ip$ for incomplete primary, cp for complete primary, is for the first cycle of secondary, cs for the second cycle of secondary, ih for incomplete higher, and ch for complete higher. DUR_i is the duration in years of the ith level of schooling—$i = p$ for primary, $s1$ for the first cycle of secondary, $s2$ for the second cycle of secondary, and h for higher.

We took account of the significant variation across countries in the standard number of years of schooling at each level. Appendix Table A.3 shows the typical duration of primary and the two levels of secondary education for each country in 1965 (derived from issues of UNESCO, *Statistical Yearbook*). We used these values to measure duration of schooling in each country and thereby neglected variations of this duration over time within a country. The data shown for 1970 in Table A.3 suggest that this assumption may not be a large source of error. When data on the duration of each cycle of secondary school were not available, we assigned one-half of the total duration of secondary school to the first cycle.

Equation (9) shows that we assigned half of the duration of primary school to persons who entered primary school but did not complete it. For secondary school, we assigned the full duration of the first cycle to everyone who entered this cycle. Then we

added the duration of the second cycle for people who entered this cycle. For higher education, we used a duration of four years for all countries, and we assigned two years to persons who entered higher school but did not complete it.

It would also be possible to construct measures of overall education capital that weighed the attainment at various levels of schooling by estimated rates of return at each level. However, the available estimates of returns to schooling, surveyed by Psacharopoulos (1985), have problems. First, the estimates tend to overstate returns because of the likely positive correlation between schooling and unmeasured characteristics such as ability. Second, the calculations are based on market earnings and therefore do not count either non-market returns or spillover benefits. We plan to consider in future research the possibilities for getting better measures of aggregate education capital.

Table 6 shows the levels of educational attainment for groups of countries that have complete data on educational attainment from 1960 to 1985. The table shows a breakdown into three major groups: developing countries (73 countries), OECD (23), and centrally planned economies (10). The developing countries are then classified into five regions: Middle East and North Africa (12), Sub-Saharan Africa (21), Latin America and the Caribbean (23), East Asia and the Pacific (10), and South Asia (7). The averages for each group were computed by weighting the attainment figures by the population aged 25 and over in each country.

Some of the principal observations from Table 6 are the following:

- For developing countries as a whole, the average of school attainment doubled from 1.8 years in 1960 to 3.6 years in 1985, while that in the OECD grew by only 30% from 6.7 years to 8.9 years. Nevertheless, the value for the OECD in 1985 (8.9 years) was still more than double that in the developing countries (3.6 years). For developing countries as a whole,

about half of the adult population in 1985 had received no formal education, whereas in the OECD, over 40 percent of the adult population had obtained some secondary education.

•There are considerable variations in profiles of educational attainment among the developing regions. In Middle East/North Africa, Sub-Saharan Africa, and South Asia, the most common characteristic is no schooling, whereas in Latin America/the Caribbean it is primary schooling. In East Asia/the Pacific, there has been a substantial change in the profile of educational attainment: the most common trait shifted from no schooling to primary schooling.

•Sub-Saharan Africa showed the lowest absolute increase in human capital over the period, from 1.5 years in 1960 to 2.7 years in 1985. Although it has achieved sizable educational attainment at the primary and secondary levels, attainment in total has been held back especially because of high dropout rates.

•Latin America had the most educational attainment among the developing countries until 1975. But East Asia has become the leader since 1980.

•The formerly centrally planned economies have the highest educational attainment. These high levels reflect especially the long history of compulsory primary education.

IV. Comparisons with Alternative Estimates

Our data set provides estimates of educational attainment at various levels of schooling, as well as figures on overall years of attainment at all levels combined. We can compare our estimates of overall years of attainment with those provided by Psacharopoulos and Ariagada (1986a); Lau, Jamison, and Louat (1991); and Kyriacou

(1991). The number of observations from 1960 to 1985 are 687 for our data set (6 observations for 106 countries, plus 51 for the 23 countries that have incomplete data), 129 for Psacharopoulos, 408 for Lau, and 500 for Kyriacou.

The correlations of our values with the others for common cells are 0.93 (111 observations) with Psacharopoulos, 0.76 with Lau (328 observations), and 0.89 with Kyriacou (424 observations). The correlation of our estimates with those of Psacharopoulos is high in every regional sub-sample, ranging from 0.76 in Sub-Saharan Africa to 0.92 in the OECD. The comparison with Lau and Kyriacou is more dependent on the region. The correlation with Lau is low for North Africa/Middle East (0.37) and East Asia/the Pacific (0.45), but high for South Asia (0.94) and the OECD (0.90). The correlations with Kyriacou range from 0.64 for the OECD to 0.89 for North Africa/Middle East.

V. Educational Attainment of the Female Population

We have also constructed a full data set on educational attainment of the female population aged 25 and over. We used the same sources and methods that we described before for educational attainment of the total adult population. We briefly sketch the data and procedures as they apply to the female population.

The basic data on female educational attainment have been collected from the census/survey data that we already discussed. Appendix Table A.4 shows the data on educational attainment of the female population aged 25 and over for each country as classified again by six levels of education. The data availability is roughly the same at that for the total population: we have lost only 12 observations (9 observations from 1960 to 1985) without the loss of any country.

We again used information on adult illiteracy to estimate the no-schooling percentage in some countries; in this case, 125 observations were available. We also

filled in the missing observations at the broad four-level classification—no schooling, first level total, second level total, and higher—by means of the same perpetual inventory method that we used for the total population.

We carried out an accuracy test by comparing the forward- and backward-flow estimates with forecasts derived from simple linear trends in the 25 countries that had complete survey data for 1960, 1970, and 1980. Because the results were essentially the same as those for the total population, we used the same procedure to fill–in the missing observations.[13]

We filled in missing observations for the sub-categories of the broad education levels—first, second, and higher—by estimating completion ratios for each country based on that country's prior or subsequent ratio and on regional dummies. The procedure is essentially the same as that used before for the total adult population. Table 7 summarizes the completion ratios for female primary and secondary education that are available from the censuses and surveys (reported in Appendix Table A.4). Table 5 contains the completion ratios for female higher education.

Finally, we constructed the full data set on educational attainment of the female population over age 25. Appendix Table A.5 presents the data set for 129 countries, including 106 countries that have complete data from 1960 to 1985. We used the data from Appendix Table A.5 to construct female human capital stocks by region. These values, contained in Table 8, show that the profiles of female educational attainment among the regions are in most cases similar to those shown in Table 6 for the total adult population.

We now compare some aspects of overall years of educational attainment for females with those of males. Table 9 summarizes the trends of female and male

---

[13]The statistics for goodness-of-fit for female attainment in the 25 countries was similar to that shown in Table 3 for total attainment in 30 countries.

attainment for the period from 1960 to 1985 by groups of countries. Some principal observations are as follows:

- In OECD countries, the difference between male and female overall years of attainment has been small: the attainment of the female population has been about 95 percent of that of the male population from 1960 to 1985. In contrast, in the developing countries, the difference has been substantial, as males had about twice as many years of schooling as females until the 1970s. Although the gap declined in the 1980s, females still had only 67 percent of male attainment in 1985.
- The gender ratio varied substantially across the developing regions. The gap has been most striking in South Asia, in which female educational attainment was only 28 percent of male attainment in 1960, but rose to 48 percent in 1985. In Middle East/North Africa and in Sub-Saharan Africa, the ratio of female to male attainment has been roughly stable at around 55 percent. In contrast, the gap has been considerably smaller in Latin America/Caribbean: female attainment was 81 percent of the male level in 1960 and 91 percent of the male level in 1985.
- The disparity by gender in education has been rapidly declining in the East Asian/Pacific countries: female attainment increased sharply from 49 percent of the male level in 1960 to 85 percent in 1985.
- Finally, in the centrally-planned economies, the ratio of female to male educational attainment rose slowly from 85% in 1960 to 91% in 1985.

VI. Concluding observations

We anticipate that this improved data set on educational attainment will have

many uses for empirical studies of economic development. It will be possible to use a broad set of panel data to examine the effects and determinants of the overall years of school attainment as well as of the composition of attainment by various levels of education. It will also be possible to assess the different influences of male and female human capital.

Our initial use of the data involves a panel study of the determinants of economic growth, physical investment, investment in human capital, and fertility. The results that we have thus far are highly preliminary, but suggest that the measures of educational attainment have considerable explanatory power.

For growth rates of real per capita GDP, the overall years of male and female school attainment at a prior date are each positive influences, but male attainment appears to be more important. For the ratio of physical investment spending to GDP, male and female attainment are again both positive influences, but male attainment is more than twice as important. In contrast, in regressions for the total fertility rate, female school attainment is negative and highly significant, whereas male attainment is insignificant. For secondary and tertiary school enrollment ratios, female educational attainment is positive and significant, but male attainment is insignificant.

The idea that the value of women's time—and hence, female educational attainment—is a key negative influence on fertility is well known; see, for example, Behrman (1990) and Schultz (1989). Moreover, a related proposition is that the higher value of women's time leads to a substitution of child quality for child quantity, that is, a higher level of female educational attainment stimulates the acquisition of human capital by children (see De Tray [1973] and Becker and Lewis [1973]). But, as far as we know, the lack of satisfactory international data has prevented adequate tests of these hypotheses for a broad panel of countries.[14] The preliminary findings suggest that our

[14]Gill and Bhalla (1992) and Bhalla and Gill (1992) have some preliminary findings that

new data set will provide strong support for the hypothesis that increased education of women leads to a sharp fall in fertility, and hence, in population growth, and to an increase in the educational attainment of children.[15] It also appears, however, that male educational attainment is more important in terms of the direct effects on GDP growth and non-human investment. This last finding likely reflects the greater labor-force role of males in most developing countries.

---

use the data constructed by Lau, Jamison, and Louat (1991).

[15]Our results about female human capital accord in some respects with the viewpoints of Summers (1992). He goes quite far, however, and even argues "... the education of girls may well be the highest return investment available in the developing world." It is unclear how to reconcile this conclusion with the findings of De Long and Summers (1992), who seem to believe that investment in machinery is the key element in economic growth. Perhaps the true key is to have educated women working with machines.

References

Barro, R.J., "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, May 1991, 407–443.

Becker, G.S. and H.G. Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, 81, March/April 1973, part II, S279–S288.

Becker, G.S., K. Murphy, and R. Tamura, "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, 98, October 1990, part 2, S12–S37.

Behrman, J.R., "Women's Schooling and Nonmarket Productivity: a Survey and a Reappraisal," unpublished, The World Bank, 1990.

Bhalla, S.S. and I.S. Gill, "Social Expenditure Policies and Welfare Achievement in Developing Countries," unpublished, State University of New York at Buffalo, December 1992.

Carceles, C., "World Literacy Prospects at the Turn of the Century: Is the Objective of Literacy for All by Year 2000 Statistically Plausible?," *Comparative Education Review*, 34, 1990, 4–20.

Chapman, D.W. and R.A. Boothroyd, "Threats to Data Quality in Developing Country Settings," *Comparative Education Review*, 32, 1988, 416–429.

De Long, J.B. and L.H. Summers, "How Robust is the Growth-Machinery Nexus?" unpublished, The World Bank, June 1992.

De Tray, D.N., "Child Quality and the Demand for Children," *Journal of Political Economy*, 81, March/April 1973, part II, S70–S95.

Fredriksen, B., "On the Collection of Education Statistics in Developing Countries: Purpose, Principles, and Problems," *International Journal of Educational Development*, 3, 1983, 291–304.

Fredriksen, B., *An Introduction to the Analysis of Student Enrollment and Flow Statistics*, Report no. PHREE/91/39, Population and Human Resources Department, The World Bank, 1991.

Gill, I.S. and S.S. Bhalla, "Externalities in New Growth Theory: the Importance of Female Human Capital," unpublished, State University of New York at Buffalo, 1992.

Kaneko, M., *The Educational Composition of the World Population: a Database*, Report no. EDT 29, Education and Training Department, The World Bank, 1986.

Kyriacou, G.A., "Level and Growth Effects of Human Capital: a Cross-Country Study of the Convergence Hypothesis," unpublished, New York University, June 1991.

Lau, L., D. Jamison, and F. Louat, *Education and Productivity in Developing Countries: an Aggregate Production Function Approach*, Report no. WPS 612, The World Bank, March 1991.

Lucas, R.E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, July 1988, 3–42.

Mulligan, C. and X. Sala–i–Martin, "Two Capital Goods Models of Economic Growth," unpublished, Yale University, 1992.

Pindyck, R.S: and D.L. Rubinfeld, *Econometric Models and Economic Forecasts*, 2nd edition, McGraw Hill, New York, 1981.

Psacharopoulos, G., "Returns to Education: a Further International Update and Implications," *Journal of Human Resources*, 20, 1985, 583–597.

Psacharopoulos, G. and A.M. Ariagada, "The Educational Composition of the Labor Force: an International Comparison," *International Labor Review*, 125, September/October 1986a, 561–574.

Psacharopoulos, G. and A.M. Ariagada, *The Educational Attainment of the Labor Force: an International Comparison*, Report no. EDT 38, Education and Training Department, The World Bank, 1986b.

Psacharopoulos, G. and A.M. Ariagada, "The International Composition of the Labor Force: an International Update," unpublished, The World Bank, 1991.

Rebelo, S., "Long Run Policy Analysis and Long Run Growth," *Journal of Political Economy*, 99, June 1991, 500–521.

Romer, P.M., "Human Capital and Growth: Theory and Evidence," *Carnegie-Rochester Conference Series on Public Policy*, 32, Spring 1990, 251–286.

Schultz, T.P., "Returns to Women's Education," PHRWD Background Paper No. PHRWD/89/001, The World Bank, 1989.

Summers, L.H., "Investing in All the People, Educating Women in Developing Countries," presented at the 1992 World Bank annual meetings, 1992.

Tilak, B.G., "Education and its Relation to Economic Growth, Poverty, and Income Distribution: Past Evidence and Further Analysis," discussion paper no. 46, The World Bank, 1989.

United Nations, *Global Estimates and Projections of Population by Sex and Age, the 1990 Assessment*, New York, 1991.

United Nations Development Programme, *Human Development Report 1990*, Oxford University Press, Oxford, 1990.

UNESCO, *Toward a Methodology for Projecting Rates of Literacy and Educational Attainment*, Division of Statistics on Education, UNESCO, Paris, 1978.

UNESCO, *Statistics of Educational Attainment and Illiteracy, 1970–1980*, Division of Statistics on Education, UNESCO, Paris, 1983.

Table 1

Breakdown of Countries by Numbers of
Census–Survey Observations for Total Population

Number of observations since 1945	Number of countries
1	22
2	30
3	39
4	29
5	6
6	2
7	1
Total	129

Note: The data refer to census-survey observations for educational attainment for the total and female populations aged 25 and over.

Table 2

Breakdown of Countries by Census-Survey Year

Census-survey year (to nearest 5-year value)	Number of countries
1945	5
1950	33
1955	15
1960	71
1965	30
1970	74
1975	44
1980	78
1985	14
Total	364
Total 1960-85	311

Note: See Table 1.

Table 3

Statistics for Ex Post Simulations for 30 countries

	No School	Primary	Secondary	Higher
1960				
Actual Mean	30.7	55.6	11.7	2.1
RMS error	11.2	17.5	13.3	2.0
Theil U	0.14	0.16	0.36	0.25
1970				
Actual Mean	26.4	52.4	17.5	3.7
RMS error	6.9	10.5	8.6	0.9
Theil U	0.11	0.09	0.19	0.10
1980				
Actual Mean	20.0	48.9	25.0	6.0
RMS error	9.7	16.5	7.6	1.9
Theil U	0.18	0.15	0.14	0.14

Notes: The actual means refer to educational attainment percentages for each category. The estimates for 1980 come from the benchmark values of educational attainment for 1970 and the forward flow computed from enrollment rates. For 1970, the estimates are a linear combination of the forward-flow value and the linear interpolation of the values from 1960 and 1980 (see the text). The estimates for 1960 come from the benchmark values of educational attainment for 1970 and the backward flow computed from enrollment rates. RMS is the root-mean-squared deviation between the actual and estimated values. Theil U is the U statistic for goodness of fit.

Table 4

Completion Ratios for Primary and Secondary Education for Total Population over Age 25, Summary by Region

I. Primary education

Region (number of countries)	Number of Observations	Average	Standard Deviation
Middle East/ North Africa (15)	14	0.38	0.23
Sub-Saharan Africa (31)	31	0.27	0.18
Latin America/ Caribbean (29)	54	0.32	0.16
East Asia/ Pacific (14)	26	0.44	0.22
South Asia (7)	6	0.31	0.28
OECD (23)	22	0.55	0.18
Centrally Planned Economies (10)	12	0.42	0.25
Total (129)	165	0.37	0.21

II. Secondary education

Middle East/ North Africa (15)	10	0.49	0.12
Sub-Saharan Africa (31)	20	0.28	0.14
Latin America/ Caribbean (29)	31	0.42	0.14
East Asia/ Pacific (14)	19	0.42	0.14
South Asia (7)	6	0.37	0.12
OECD (23)	26	0.43	0.22
Centrally Planned Economies (10)	4	0.39	0.18
Total (129) .	116	0.40	0.17

Note: The completion ratio for primary education is the ratio of the attainment percentage for complete primary to the attainment percentage for total primary from Appendix Table A.1. The completion ratio for secondary education is the ratio of the attainment percentage for the second cycle to the attainment percentage for total secondary from Appendix Table A.1.

Table 5

Completion Ratios for Higher Education

Country	Year	Higher educational attainment (percent of population aged 25 and over) and completion ratio			
		Total		Female	
Middle East and North Africa					
Bahrain	1982	6.5	0.47	4.1	0.45
Egypt	1976	3.8	0.88	1.6	0.81
Israel	1983	23.9	0.47	22.0	0.41
Jordan	1979	6.7	0.61	3.3	0.36
Tunisia	1975	2.5	0.90	0.5	1.00
Regional average			0.67		0.60
Sub-Saharan Africa					
Cameroon	1976	0.9	0.66	0.3	0.67
Lesotho	1976	0.1	1.00	0.1	1.00
Liberia	1974	1.4	0.74	0.8	0.75
Malawi	1977	0.3	1.00	0.1	1.00
Mauritius	1983	3.6	0.42	1.9	0.32
Rwanda	1978	0.2	0.88	0.0	--
Regional average			0.78		0.75
Latin American and the Caribbean					
Argentina	1980	7.3	0.60	6.2	0.64
Brazil	1980	5.1	0.68	4.3	0.64
Haiti	1982	1.5	0.75	1.2	0.67
Mexico	1980	9.2	0.59	--	--
Paraguay	1982	4.2	0.79	4.1	0.84
Regional average			0.68		0.70
East Asia and the Pacific					
Hong Kong	1981	7.1	0.57	4.9	0.47
Indonesia	1980	1.4	0.44	0.7	0.46
Korea	1980	9.0	0.74	4.1	0.74
Malaysia	1980	1.9	0.85	1.1	0.81
Philippines	1980	15.3	0.65	15.1	0.70
Taiwan	1980	9.3	0.51	5.5	0.44
Thailand	1980	2.9	1.00	2.3	0.99
Regional average			0.68		0.66

Table 5, continued

Country	Year	Higher educational attainment (percent of population aged 25 and over) and completion ratio			
		Total		Female	
South Asia					
Pakistan	1981	2.1	0.86	0.8	0.87
Sri Lanka	1981	2.1	0.54	1.9	0.46
Regional average			0.70		0.66
OECD					
Belgium	1977	9.4	0.81	7.7	0.69
Canada	1981	40.2	0.27	37.5	0.22
Finland	1980	9.0	0.58	8.4	0.51
France	1982	8.6	0.51	7.7	0.38
Greece	1981	7.6	0.96	4.9	0.98
Japan	1980	14.4	0.62	9.6	0.31
New Zealand	1981	27.7	0.38	24.8	0.37
Norway	1980	12.0	0.44	8.9	0.32
Portugal	1981	5.4	0.48	4.9	0.34
United States	1980	31.8	0.60	28.1	0.56
Regional average			0.57		0.47
Centrally Planned Economies					
Cuba	1981	9.8	0.76	9.1	0.78
Hungary	1980	7.0	0.96	5.1	0.96
Regional average			0.86		0.87

Source: Constructed from Kaneko (1986).

Note: The completion ratio for higher education is the ratio of the attainment percentage for complete tertiary schooling, which includes university graduates with degrees and post-graduates, to the attainment percentage for total tertiary schooling, which includes all persons who entered tertiary school.

Table 6

Trends of Educational Attainment by Region

Region/ Group	Year	Pop. over age 25 (millions)	highest level attained (percentage of the population over 25) No School	Primary total (complete)	Secondary total (complete)	Higher total (complete)	Avg. years of school
All	1960	468	68.4	25.8 (8.3)	5.0 (1.9)	0.8 (0.5)	1.76
Developing	1965	524	65.1	28.2 (9.8)	5.6 (2.2)	1.2 (0.8)	2.01
Countries	1970	585	61.0	30.2 (11.0)	7.0 (2.8)	1.7 (1.2)	2.36
(73	1975	658	57.3	30.9 (9.8)	9.3 (3.6)	2.5 (1.7)	2.71
countries)	1980	753	54.9	28.9 (8.7)	13.0 (5.1)	3.2 (2.2)	3.10
	1985	872	49.7	31.4 (10.1)	14.6 (5.9)	4.4 (3.0)	3.56
Middle East/	1960	20	84.2	11.2 (3.8)	3.5 (1.7)	1.0 (0.6)	1.02
North Africa	1965	22	81.8	12.1 (4.6)	4.6 (2.3)	1.4 (0.8)	1.24
(12	1970	25	76.4	15.6 (5.6)	6.2 (3.1)	1.8 (1.1)	1.60
countries)	1975	29	68.9	18.9 (6.6)	9.6 (4.9)	2.6 (1.6)	2.21
	1980	35	61.6	22.9 (7.9)	11.7 (6.2)	3.8 (2.3)	2.77
	1985	43	52.8	26.5 (9.1)	16.0 (8.6)	4.8 (3.0)	3.51
Sub-Saharan	1960	40	74.5	19.1 (6.0)	5.9 (1.5)	0.5 (0.4)	1.48
Africa	1965	45	70.6	22.6 (5.7)	6.1 (1.5)	0.7 (0.5)	1.62
(21	1970	51	66.7	25.3 (5.4)	7.0 (1.7)	1.0 (0.8)	1.85
countries)	1975	58	62.5	29.3 (6.0)	7.2 (1.5)	1.0 (0.8)	2.00
	1980	66	55.6	35.1 (6.7)	8.6 (1.5)	0.8 (0.6)	2.31
	1985	77	48.1	41.7 (8.5)	9.3 (1.6)	1.0 (0.8)	2.67
Latin	1960	82	41.6	47.0 (13.0)	9.5 (3.9)	1.8 (1.2)	3.01
America/	1965	92	38.4	50.1 (15.0)	9.4 (4.0)	2.1 (1.3)	3.17
Caribbean	1970	104	34.7	52.2 (17.9)	10.8 (4.5)	2.5 (1.6)	3.50
(23	1975	120	29.6	55.6 (12.4)	10.7 (4.3)	4.1 (2.7)	3.67
countries)	1980	139	28.4	54.0 (13.6)	12.4 (4.9)	5.4 (3.5)	4.01
	1985	162	22.4	56.6 (13.2)	13.9 (5.5)	7.1 (4.6)	4.47
East Asia/	1960	79	61.2	31.6 (15.2)	5.7 (2.3)	1.6 (1.1)	2.26
Pacific	1965	89	53.5	36.7 (17.1)	7.7 (3.3)	2.1 (1.4)	2.75
(10	1970	100	42.1	45.3 (19.9)	9.8 (4.3)	2.8 (1.9)	3.41
countries)	1975	113	36.9	47.3 (19.1)	12.3 (5.7)	3.6 (2.4)	3.83
	1980	131	30.1	49.4 (17.4)	15.6 (7.6)	4.8 (3.2)	4.38
	1985	154	23.6	51.3 (22.9)	18.8 (9.4)	6.3 (4.3)	5.19
South Asia	1960	248	77.3	19.3 (5.2)	3.2 (1.2)	0.1 (0.1)	1.30
(7	1965	275	75.5	20.2 (6.9)	3.6 (1.3)	0.6 (0.5)	1.51
countries)	1970	304	74.0	19.8 (7.2)	4.9 (1.9)	1.2 (0.9)	1.77
	1975	338	71.9	18.1 (6.6)	8.1 (3.0)	1.9 (1.4)	2.17
	1980	382	72.2	12.3 (4.4)	13.2 (4.9)	2.3 (1.6)	2.49
	1985	436	69.0	13.7 (4.8)	14.1 (5.3)	3.2 (2.3)	2.81

Table 6, continued

Region/ Group	Year	Pop. over age 25 (millions)	highest level attained (percentage of the population over 25) No School	Primary total (complete)	Secondary total (complete)	Higher total (complete)	Avg. years of school
OECD (23 countries)	1960	362	6.4	61.0 (33.8)	25.5 (9.8)	7.0 (4.1)	6.71
	1965	383	6.0	58.0 (33.7)	27.9 (11.5)	8.2 (4.8)	7.03
	1970	404	5.2	54.0 (31.4)	31.3 (13.9)	9.5 (5.6)	7.42
	1975	435	5.4	47.7 (25.3)	34.2 (16.5)	12.8 (7.3)	7.88
	1980	467	4.6	39.4 (19.9)	40.2 (22.5)	15.9 (9.1)	8.65
	1985	501	3.3	37.7 (18.3)	40.8 (20.1)	18.2 (10.5)	8.88
Centrally Planned Economies (10 countries)	1960	183	5.0	68.9 (26.0)	22.3 (9.0)	3.9 (3.4)	6.83
	1965	202	5.3	62.1 (25.7)	27.6 (10.9)	5.0 (4.3)	7.29
	1970	208	4.0	53.4 (22.7)	36.3 (14.3)	6.4 (5.5)	7.97
	1975	221	3.7	47.9 (20.2)	40.9 (16.1)	7.5 (6.5)	8.33
	1980	237	2.7	39.4 (17.0)	49.9 (12.3)	8.0 (6.9)	8.78
	1985	253	2.3	36.1 (14.3)	51.9 (20.4)	9.8 (8.4)	9.17

Note: The regional averages are weighted by each country's population aged 25 and over. Each group/region contains the countries listed in Appendix Table A.2 (but only those that have complete data on educational attainment from 1960 to 1985).

Table 7

Completion Ratios for Primary and Secondary Education for Female Population over Age 25, Summary by Region

I. Primary education

Region (number of countries)	Number of Observations	Average	Standard Deviation
Middle East/ North Africa (15)	14	0.38	0.21
Sub-Saharan Africa (31)	31	0.28	0.23
Latin America/ Caribbean (29)	51	0.32	0.17
East Asia/ Pacific (14)	26	0.41	0.23
South Asia (7)	6	0.29	0.24
OECD (23)	20	0.53	0.18
Centrally Planned Economies (10)	12	0.40	0.25
Total (129)	160	0.37	0.22

II. Secondary education

Region (number of countries)	Number of Observations	Average	Standard Deviation
Middle East/ North Africa (15)	9	0.46	0.22
Sub-Saharan Africa (31)	21	0.29	0.19
Latin America/ Caribbean (29)	28	0.43	0.16
East Asia/ Pacific (14)	19	0.38	0.18
South Asia (7)	5	0.32	0.14
OECD (23)	25	0.43	0.22
Centrally Planned Economies (10)	4	0.40	0.20
Total (129)	111	0.39	0.19

Note: The completion ratio for primary education is the ratio of the attainment percentage for complete primary to the attainment percentage for total primary from Appendix Table A.4. The completion ratio for secondary education is the ratio of the attainment percentage for the second cycle to the attainment percentage for total secondary from Appendix Table A.4.

Table 8

Trends of Female Educational Attainment by Region

Region/ Group	Year	Pop. over age 25 (millions)	highest level attained (percentage of the population over 25) No School	Primary total (complete)	Secondary total (complete)	Higher total (complete)	Avg. years of school
All Developing Countries (73 countries)	1960	230	78.2	18.0 (5.8)	3.3 (1.3)	0.4 (0.3)	1.16
	1965	258	75.3	20.4 (6.7)	3.8 (1.5)	0.6 (0.4)	1.34
	1970	290	71.3	23.2 (7.9)	4.5 (1.8)	0.9 (0.6)	1.59
	1975	327	67.7	25.0 (6.9)	5.9 (2.3)	1.4 (0.9)	1.84
	1980	374	63.7	25.9 (6.8)	8.5 (3.3)	2.0 (1.4)	2.22
	1985	434	54.4	32.6 (9.7)	10.1 (4.2)	2.9 (1.9)	2.85
Middle East/ North Africa (12 countries)	1960	10	89.0	7.9 (2.8)	2.6 (1.2)	0.5 (0.3)	0.72
	1965	11	87.7	8.2 (3.0)	3.4 (1.6)	0.7 (0.4)	0.84
	1970	13	86.1	8.9 (3.2)	4.0 (1.9)	1.1 (0.5)	0.98
	1975	15	79.2	13.5 (4.7)	5.7 (2.8)	1.6 (0.8)	1.42
	1980	18	73.0	17.3 (5.9)	7.2 (3.6)	2.4 (1.2)	1.85
	1985	21	64.1	22.2 (7.5)	10.6 (5.3)	3.1 (1.6)	2.50
Sub-Saharan Africa (21 countries)	1960	21	81.9	12.6 (4.1)	5.3 (1.3)	0.3 (0.2)	1.12
	1965	23	79.6	14.9 (2.3)	5.2 (1.3)	0.4 (0.3)	1.15
	1970	27	76.7	16.9 (2.6)	5.7 (1.3)	0.7 (0.5)	1.32
	1975	30	72.6	20.8 (1.9)	6.0 (1.1)	0.7 (0.5)	1.43
	1980	34	65.3	27.5 (1.8)	6.5 (1.0)	0.6 (0.5)	1.69
	1985	40	59.3	32.7 (1.6)	7.4 (0.9)	0.5 (0.4)	1.92
Latin America/ Caribbean (23 countries)	1960	41	46.3	44.0 (12.8)	8.8 (3.9)	0.9 (0.6)	2.71
	1965	47	43.0	47.4 (14.1)	8.6 (3.9)	1.1 (0.7)	2.83
	1970	53	38.3	50.6 (16.5)	9.7 (4.5)	1.3 (0.9)	3.14
	1975	61	32.4	54.6 (11.2)	10.2 (5.0)	2.8 (1.8)	3.38
	1980	71	29.0	55.0 (12.3)	12.0 (6.0)	3.9 (2.6)	3.79
	1985	82	23.3	57.3 (11.8)	14.0 (7.1)	5.5 (3.7)	4.27
East Asia/ Pacific (10 countries)	1960	40	73.1	22.5 (10.5)	3.3 (1.3)	0.9 (0.6)	1.50
	1965	45	64.8	29.3 (12.7)	4.6 (1.8)	1.3 (0.9)	1.94
	1970	50	53.2	38.8 (16.1)	6.3 (2.6)	1.7 (1.2)	2.58
	1975	57	47.4	41.7 (16.0)	8.5 (3.8)	2.4 (1.7)	2.99
	1980	66	39.7	45.3 (15.2)	11.6 (5.2)	3.6 (2.5)	3.56
	1985	78	23.7	56.7 (25.2)	14.9 (7.4)	4.7 (3.3)	4.76
South Asia (7 countries)	1960	118	89.5	9.2 (2.4)	1.2 (0.4)	0.0 (0.0)	0.55
	1965	132	88.5	9.9 (3.1)	1.6 (0.5)	0.2 (0.2)	0.69
	1970	147	87.2	10.4 (3.4)	1.9 (0.6)	0.4 (0.3)	0.79
	1975	164	85.8	10.1 (3.3)	3.4 (1.1)	0.6 (0.4)	0.99
	1980	185	84.3	8.4 (2.6)	6.5 (2.0)	1.0 (0.6)	1.27
	1985	212	76.0	15.1 (4.8)	7.4 (2.3)	1.6 (1.0)	1.81

Table 8, continued

Region/ Group	Year	Pop. over age 25 (millions)	highest level attained (percentage of the population over 25) No School	Primary total (complete)	Secondary total (complete)	Higher total (complete)	Avg. years of school
OECD (23 countries)	1960	190	7.5	60.9 (33.1)	26.2 (10.3)	5.5 (2.8)	6.53
	1965	201	7.1	58.1 (33.1)	28.9 (12.4)	6.3 (3.2)	6.85
	1970	212	6.2	54.8 (31.7)	32.1 (14.9)	7.1 (3.6)	7.19
	1975	228	6.3	48.4 (25.4)	35.5 (17.9)	10.2 (5.1)	7.67
	1980	245	5.5	40.4 (20.2)	40.2 (23.0)	13.9 (6.7)	8.39
	1985	262	4.0	39.0 (18.6)	40.9 (20.7)	16.1 (7.9)	8.61
Centrally Planned Economies (10 countries)	1960	105	6.3	71.5 (23.5)	19.4 (8.1)	2.8 (2.4)	6.36
	1965	114	6.0	66.2 (23.8)	24.1 (9.7)	3.8 (3.3)	6.81
	1970	117	4.8	57.6 (21.5)	32.9 (13.2)	5.1 (4.4)	7.53
	1975	123	4.3	52.3 (19.3)	37.3 (15.1)	6.3 (5.5)	7.91
	1980	131	3.5	43.4 (16.3)	46.1 (12.0)	7.0 (6.1)	8.37
	1985	139	3.0	39.2 (13.3)	49.0 (19.8)	8.7 (7.6)	8.79

Note: The regional averages are weighted by each country's female population aged 25 and over. Each group/region contains the countries listed in Appendix Table A.5 (but only those that have complete data on educational attainment from 1960 to 1985).

Table 9
Trends of Gender Ratio in Educational Attainment by Region

Region/Group	year	Female		Male		Gender ratio*
		Pop. (millions)	Avg. years of school	Pop. (millions)	Avg. years of school	
All Developing Countries (73 countries)	1960	230	1.16	238	2.34	49.5
	1965	258	1.34	265	2.67	50.2
	1970	290	1.59	295	3.12	50.9
	1975	327	1.84	331	3.58	51.5
	1980	374	2.22	379	3.96	56.1
	1985	434	2.85	439	4.26	67.0
Middle East/ North Africa (12 countries)	1960	10	0.72	10	1.31	54.5
	1965	11	0.84	11	1.64	50.9
	1970	13	0.98	13	2.24	43.6
	1975	15	1.42	14	3.02	47.1
	1980	18	1.85	18	3.69	50.3
	1985	21	2.50	21	4.52	55.3
Sub-Saharan Africa (21 countries)	1960	21	1.12	19	1.86	60.1
	1965	23	1.15	22	2.13	54.1
	1970	27	1.32	25	2.41	54.7
	1975	30	1.43	28	2.60	55.0
	1980	34	1.69	32	2.96	57.1
	1985	40	1.92	37	3.47	55.4
Latin America/ Caribbean (23 countries)	1960	41	2.71	40	3.32	81.4
	1965	47	2.83	45	3.52	80.5
	1970	53	3.14	51	3.87	81.3
	1975	61	3.38	59	3.97	85.1
	1980	71	3.79	68	4.23	89.7
	1985	82	4.27	79	4.69	91.1
East Asia/ Pacific (10 countries)	1960	40	1.50	39	3.04	49.2
	1965	45	1.94	44	3.58	54.1
	1970	50	2.58	50	4.27	60.3
	1975	57	2.99	56	4.69	63.8
	1980	66	3.56	65	5.23	68.0
	1985	78	4.76	76	5.63	84.7
South Asia (7 countries)	1960	118	0.55	129	1.98	28.0
	1965	132	0.69	142	2.28	30.1
	1970	147	0.79	157	2.69	29.3
	1975	164	0.99	174	3.29	30.0
	1980	185	1.27	197	3.63	35.0
	1985	212	1.81	224	3.75	48.2

Table 9, continued

Region/Group	year	Female		Male		Gender ratio*
		Pop. (millions)	Avg. years of school	Pop. (millions)	Avg. years of school	
OECD	1960	190	6.53	172	6.89	94.8
(23 countries)	1965	201	6.85	182	7.22	94.9
	1970	212	7.19	192	7.67	93.7
	1975	228	7.67	207	8.11	94.6
	1980	245	8.39	222	8.95	93.8
	1985	262	8.61	239	9.16	94.0
Centrally	1960	105	6.36	78	7.46	85.2
Planned	1965	114	6.81	88	7.92	86.0
Economies	1970	117	7.53	91	8.53	88.3
(10 countries)	1975	123	7.91	98	8.87	89.1
	1980	131	8.37	106	9.28	90.2
	1985	139	8.79	114	9.63	91.3

*The gender ratio is defined as the ratio of female to male average years of schooling, multiplied by 100.

Note: Pop. refers to the population aged 25 and over.