GENDER DIFFERENCES IN RETURNS TO SCHOOLING:

AN INTERNATIONAL CROSS-COUNTRY STUDY

PETER THIEL

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF ARTS

Department of Economics

ILLINOIS STATE UNIVERSITY

1995

UMI Number: 9603525

DISSERTATION APPROVED:

July 26 95 Rati Ram

Date Rati Ram, Chair

7/26/95 Anthony L. Ostrosky

Date Anthony L. Ostrosky

26-Aug-95 Mark S. Walbert

Date Mark S. Walbert

GENDER DIFFERENCES IN RETURNS TO SCHOOLING:

AN INTERNATIONAL CROSS-COUNTRY STUDY

Peter Thiel

65 Pages August 1995

This research seeks to determine the impact of female education on income relative to that of male education. The study addresses the question: does investment in female education yield a greater rate of return in the form of higher flows of income, than investment in male education?

Multiyear multicountry data on output per capita and mean years of schooling per male and female were used in the estimation of a modified version of the simple Mincerian ''schooling model''. The purpose of using a cross-country sample was to estimate the relative impact of female and male education on real output in a global context as well as in low-income, middle-income and high-income groups of countries. 'Intercept' dummy variables for individual countries were used to control for some of the country-specific characteristics of the sample observations. The sample consisted of 28 years of observations for each of 82 countries.

The results indicate that the rate of return to schooling is larger for females than for males. This is true for the entire sample and also for the three income-

differentiated subsamples. With regard to the patterns between the subsamples, the results indicate that the excess of the female rate of return over the male rate of return might decrease as countries evolve from low income levels to middle income levels, but seems to start rising again as countries further progress to high income levels.

APPROVED:

July 26, 95 — Rati Ram
Date — Rati Ram, Chair

7/26/95 — Anthony L. Ostrosky
Date — Anthony L. Ostrosky

26-Aug-95 — M. S. Walbert
Date — Mark S. Walbert

ACKNOWLEDGMENTS

The writer wishes to thank Dr. Rati Ram, Chair of his Advisory Committee, and the other members of the committee, Dr. Anthony Ostrosky, and Dr. Mark Walbert, for all the time, interest and assistance they so steadily provided during the writing of this dissertation. The writer also wishes to thank Dr. Elizabeth King for making the schooling dataset available.

P.T.

CONTENTS

TABLES

FIGURES

CHAPTER I

THE PROBLEM AND ITS BACKGROUND

Statement of the Problem

The idea that education is an important factor in enhancing productivity and output is not new. In fact, Schultz (1963) and Becker (1964) thought education to be such an important productive resource that they referred to it as ''human capital''. For economists, education is, to a certain extent, equivalent to machinery and equipment: the more of it that a worker has, the more productive he or she will be. Because machinery and equipment are called ''capital'' by economists, they naturally chose the term ''human capital'' to refer to education. Thus, a person spending time in school is seen by economists as making an ''investment'' in his or her human capital.

Enhancement of productivity and output is naturally related to economic growth and development. Therefore, it is logical to expect education to be an essential component of economic development and growth, and several models of economic growth have used the concept of human capital.

Given the theoretical importance that economists have given to education, considerable research has been conducted on the effect that human capital has on economic performance. Notwithstanding some problems involved with the empirical work on the role of human capital in growth, research in the area has been useful in several ways. However, there is much work that still needs to be done. For example, most of the research on returns to schooling has been conducted from intra-country data and very little cross-country research has been done in this area. Apart from providing greater sample size and variance, cross-country data could provide a direct global perspective on the rates of return. Such cross-country studies may also reveal important patterns in the rates of return across countries or regions. Also, as Ram (1996) notes, when countries, rather than individuals, are the units of observation, the worry about omission of such factors as "ability" and experience may be less troublesome.

Apart from the fact that only a limited amount of cross-country research has been conducted on the education-output nexus to estimate rates of return to schooling, it is also true that few articles have been written in the human-capital literature that assess the difference between the productivity effects of male and female schooling in a cross-country context. Even though Psacharopoulos (1973) has worked for a long time addressing this particular question, compiling and comparing data from many countries around the world, there have been very few cross-country studies that estimate educational rates of return by gender. This is so in spite of the fact that, historically, educational attainment of women in most areas of the world has been lower than that of men. Furthermore, this topic is particularly important in this late twentieth century, when developing economies are modernizing and becoming ever more inclined to take advantage of the untapped potential of their female labor force.

In many countries throughout the developing world, women's education tends to be restricted, especially at higher levels. This could be because of a tendency to consider that the resources used to educate females are wasted, since females do not tend to actively seek market work, due to their role in the household. However, many economists, such as Psacharopoulos and Woodhall (1985, pp. 93, 109 and 297), recognize that female education is a key to economic development. Besides the fact that women are a significant part of the labor force throughout the world, it is through women that children get their first and most crucial care and instruction. In addition, there is the positive effect that better education of women has on child care and nourishment, and the economic implications of this are significant. It is also well known that better educated women tend to have fewer children, being thus better able to care for them and educate them. Nevertheless, the question remains as to just how worthwhile it is to direct more effort and resources to female education. For this reason it is clear that further cross-country schooling studies that estimate the social payoff by gender are necessary.

Statement of the Purpose

The purpose of this research is to use multicountry data to estimate the impact that female education has on aggregate output or income, as indicated by the 'schooling model' developed by Mincer (1974), and to find out how different, if at all, it is from the impact of male education.

This research is an extension of the work done by Ram (1996), who adapted the Mincer (1974) schooling model to estimate returns to schooling on a cross-country basis. The main additional contribution is that this research discriminates by gender in order to gain some insights on gender-specific differentials in returns to schooling.

Review of Subsequent Chapters

The remainder of this dissertation is organized into five chapters. Chapter II dwells on the existing literature that is concerned with returns to schooling by gender. Chapter III describes Mincer's schooling model and discusses how this model is modified and used in this study. Chapter IV explains the sources of data used and the estimation methods employed. Chapter V discusses the estimation results obtained. Finally, Chapter VI provides a brief overall discussion of the dissertation and offers suggestions for future research on the topic.

CHAPTER II

PREVIOUS RESEARCH ON THE TOPIC

A Historical Perspective

Psacharopoulos (1973, p. 1) pointed out that even though the economic consequences of education became an important issue in the late nineteen fifties, its importance had been discussed much earlier, and ''in fact it was a rediscovery, since people as far back as Adam Smith and as recent as Marshall had already written about the economic consequences of education.'' For some reason, economists avoided the issue of investment in education for a very long time. T.W. Schultz (1971, p. 26) offered the explanation that moral and philosophical concerns were to blame. He argued that the very thought of investment in human beings offended some people. It gave the impression of considering persons as mere physical things, almost as property. ''For man to look upon himself as a capital good, even if it did not impair his freedom, may seem to debase him.'' Still, he had come to the conclusion that advances in technology and conventional inputs could not fully explain the observed increases in productivity and that education was likely to have a strong relevance to the issue.

Throughout the nineteen fifties, he searched for missing factors to account for the gains in productivity. From this search he determined (1971, p. v) that ''the role of acquired abilities of human agents'' was a major source of the unexplained gains in productivity.

During the late 1950s, Solow (1956, 1957) had developed macroeconomic growth models that considered the accumulation of physical capital to be the main source of economic growth. However, empirical estimation of such neoclassical growth models at the time produced high unexplained ''residuals'' which could be attributed to the omission of explanatory variables, one of which, it was thought, might be the varying skills of workers. As a result, economists started to treat education as a form of human capital.

Fabricant (1959), for example, was concerned with the measurement of productivity and with the factors that led to improvements in productivity. He argued that productivity indices at that time treated labor as being largely homogeneous, thus ignoring differences in skills, levels of education, and lengths of experience.

With this in mind, Denison (1962) worked on the development of labor quality indexes based on educational background, to account for differences in the skills and abilities of workers. He then used these indexes to help explain the economic growth rates in various industrialized countries. To develop his indexes, Denison grouped workers according to their level of formal education, and compared the differences in earnings across educational groups. He also compared workers' average level of formal education and earnings in a given year with those ten years later. He assumed that three-fifths of the registered increases in earnings across time were due to greater levels of education, the remainder being accounted for by technological advances and capital accumulation. By interpolating, he was able to create indexes on a year by year basis.

Denison's indexes, though useful, proved to be limited in important ways, the most important of which was that these indexes did not take account of the present value of the flows of all future incomes resulting from different levels of education. For this purpose, a better approach was the ''schooling model'' developed by Mincer (1974). This model was designed to study the economic benefits derived from education, namely, the returns from investment in human capital or returns to schooling. Mincer worked on the problem of empirically measuring the income effects of investment in human capital. This permitted the researchers to estimate the real rate of return to schooling. The econometric model that his efforts produced is simple to use, because it is based on a semi-log regression framework. For this reason, it is widely used for empirical studies.

More recently, economists have attempted to incorporate human capital in their growth models by extending Solow's

(1956, 1957) framework. Lucas (1988) developed a macroeconomic growth model which treated the level of skills of the labor force as an endogenous variable. In other words, his model took into account the fact that level of skills will vary as a result of changes in the economy. Lucas was also concerned with the effects that the international environment may have on an economy, and came to the conclusion that the international flows of resources may be an important factor in the performance of an economy. In other words, economists should focus more on the global economy. Another more recent study is that by Mankiw, Romer and Weil (1992). This study emphasizes the need to extend Solow's (1956, 1957) model to include human capital as well as physical, in order to predict more accurately the magnitudes, as well as the direction, of the effects of savings and population growth on an economy.

Thus, macroeconomic models through the past few decades have placed more and more emphasis on the need to account for human capital and on the need to focus more on the global economy. Good estimates of the returns to schooling, preferably from an international perspective, are useful in this context. It is therefore, necessary to engage in more empirical research to find more about returns to schooling.

Previous Empirical Studies on This Topic

It is noteworthy that through the search for literature on this topic, only one study by T. P. Schultz (1994) was found that applies the Mincer schooling model on a cross-country basis and that discriminates by gender. Even this study mentioned the gender-specific returns only briefly in the end notes of his article (pp. 53-54).

As for other research on rates of return that discriminates by gender, there have been many studies. One very important contribution on this subject has been made by Psacharopoulos (1973) who compiled and compared studies from all over the world on returns to schooling. He updated and expanded that compilation several times and the updates in Psacharopulos (1985, 1993) are particularly useful. The latest (1993) update compiles information from 82 studies for 42 countries all over the world. He then uses all of these studies to discern key world patterns on rates of return to schooling, one of which concerns the gender differential in returns to schooling. He lists summary statistics on returns to schooling for males and females, grouping them according to the level of education that the various researchers studied. It must be noted that due to the wide differences across studies and countries, the rates of return to schooling reported in the studies varied widely, with a range from 2% to 35.5%.

The largest of these groups was composed of 55 studies on 34 countries that used total schooling in their calculations, using the Mincerian method. For this specific group Psacharopoulos (pp. 14-15) found that the rate of return to schooling for males and females was 11.1% and 12.4% respectively. This suggests that, overall, rates of return to schooling are greater for females than for males. However, he found a mixed picture in the other smaller groups of studies: those that based their estimates on primary school alone averaged 20.1% for males and 12.8% for females; those based on secondary schooling averaged 13.6% and 18.4% respectively; and finally, those that concentrated on the college level yielded 13.4% and 12.7% respectively.

Psacharopoulos (1993) does an excellent job in presenting the state of the current research on this topic and in summarizing the findings. However, it provides only limited information concerning the methodologies and the problems encountered in the individual studies. It is therefore necessary for this research to look more closely into some of these studies.

For this purpose, four relatively recent articles are reviewed that use Mincer's schooling model, or a close variant, and that control for gender differences. These articles are by: Psacharopoulos & Ng (1992); McMahon (1991); Hertz & Khandker (1991); and T. P. Schultz (1994). The first three articles are thorough and extensive, where each reported results on regressions conducted on several different samples. The fourth, though not nearly so extensive, is important in the sense that it estimates a cross-country regression.

Psacharopoulos & Ng (1992) present estimates for each of 18 Latin America countries; for some of those countries two different years are covered and the estimates are computed separately for each year. McMahon (1991) presents data encompassing 20 different years for the United States. Hertz and Khandker (1991) segment their data on the Peruvian economy by region, providing separate estimates for each region, as well as for Peru as a whole. Schultz (1994) uses data on 65 countries to make two cross-country regression analyses. With respect to the treatment used in the four studies, they are similar enough to be roughly comparable. Table 1 provides a synthetic comparison of these studies.

The proxies used for income are very similar in the first three studies. The income period varies from being yearly (McMahon and Schultz), monthly (Psacharopoulos & Ng) or hourly (Hertz & Khandker) earnings, but these differences should have no major effect on the estimates. All three use earnings before taxes. Schultz, however, uses an entirely different proxy for income, due to the cross-country nature of the study. Instead of earnings before taxes, he uses GDP per adult as the proxy for income.

The proxies used for education are not quite as similar. Both Psacharopoulos & Ng and Hertz & Khandker use the total number of years spent in school combined with the years of experience and both use Mincer's schooling model that includes experience. These two are indeed working with

Table 1
A Comparison of Empirical Studies Using Mincer's Schooling Model or a Close Variant, that Control for Gender.

Studies:
1) Psacharopoulos and Ng (1992)
2) McMahon (1991)
3) Hertz and Khandker (1991)
4) Schultz (1994)

Income proxy used:
1) Individual worker's earnings per month.
2) Individual worker's net yearly earnings before taxes.
3) Individual worker's hourly wage rate.
4) GNP per adult (potential workers over age 15).

Education proxy used:
1) Individual worker's total years of schooling along with years of experience.
2) Individual worker's public and private costs of high school plus foregone earnings costs.
3) Individual worker's total years of schooling along with years of experience.
4) Expected years of enrollments for all three levels of schooling, lagged by a decade.

Scope of sample:
1) 18 countries some with 2 years of data. Separate estimates are given for each case.
2) 20 years of U.S. national data, with separate estimates for each year.
3) Peru national data on 5,100 households and 26,000 individuals.
4) 65 countries with separate regressions performed for 1970 and 1980 data.

Model Used:
1) Mincer schooling model with OLS.
2) Net present value.
3) Mincer schooling model with OLS.
4) A variant of the Mincer schooling model modified to control for gender differences and applied in a cross-country setting. Estimation with OLS.

---

the same proxy for education. Schultz, on the other hand, uses the expected years of male and female school enrollments lagged by a decade as a proxy for schooling. This particular proxy, though conveniently simple, is not the most desirable, for expected years of schooling at a point in time may not account accurately for the overall stock of educational capital of workers in the labor force. Finally, McMahon uses as a proxy for education the full public and private institutional costs of high school plus foregone earnings costs. The notable difference is that this proxy is in terms of costs, where the unit of measurement is currency, while the others use time, in years. The problem with this proxy is the difficulty of estimating the costs correctly. Even though this proxy could yield more precise estimate of investment in education, the model is more complicated to estimate than the one with simple years of schooling proxy. This problem might cause estimates to diverge from those obtained with the simpler proxy.

With respect to the actual econometric model used, Psacharopoulos & Ng and Hertz & Khandker use the Mincer schooling model that includes experience, which makes these two studies more easily comparable. The model used by Schultz naturally does not include ''experience'', and he modifies it to control for gender and applies it in a cross-country setting, where the units of observation are countries, rather than individuals. Finally, the econometric model presented by McMahon has the net present value structure. This should not create much of a problem though, because the schooling model is based on equating net present values for different years of schooling. All studies except McMahon use the ordinary least squares method of estimation. Hertz & Khandker use the maximum-likelihood method as well, but the estimates obtained with this method do not indicate the rate of return for females and males separately. McMahon calculated the rates of return directly from his data by equating the present value of future flows income for workers with different levels of education.

The nature of the samples is the major difference between the four studies. Psacharopoulos & Ng study 18 Latin American countries, with separate regressions for each country. Schultz runs two regressions, one for 1970 and another for 1980, involving 65 countries. On the other hand, Hertz & Khandker do an in-depth study of Peru alone, and McMahon works with annual data for the United States. Given that Psacharopoulos & Ng and Hertz & Khandker deal with developing economies while McMahon concentrates on a major industrialized nation, the results from these two groups may be expected to differ somewhat.

Table 2 shows the means of the estimates presented in the four studies. This table suggests that returns in the United States are larger than those in developing nations. This might appear counterintuitive, because diminishing returns are expected to apply to education. The richer the country, the more educated its work force, and so, lower returns on education should be observed. Other international

Table 2
Mean Estimates of Returns to Schooling From the Studies Presented in Table 1.

Study	Males	Females	Total
Psacharopoulos and Ng (1992)	7.56%	8.20%	7.81%
McMahon (1991)	12.14%	10.90%	11.29%
Hertz and Khandker (1991)	8.12%	9.01%	Not given
Schultz (1994)	10.40%	17.55%	Not given
Mean of all four studies	9.55%	11.41%	

studies such as Ram (1996) and Psacharopoulos (1993) strongly support such a scenario. However, we must also account for the fact that richer countries tend to have greater amounts of physical capital, the use of which may require greater levels of education in the labor force. Thus, education could yield greater returns in countries with high levels of physical capital. At any rate, the differences in the findings are probably the result of the differences in the methodology of estimation used in the different studies. In fact, the range of estimates compiled by Psacharopulos (1993) appears much larger than that in these four studies.

With regard to the overall means of the estimates, the mean for the estimated returns to schooling of males is 9.55% and that for females is 11.41%. The difference in the means between males and females suggests slightly higher returns for females than for males, but it is unclear whether this difference is significant, given the limitations imposed by the differences across studies. These averages are fairly similar to those found by Psacharopoulos (1993) of 11.1% and 12.4% for males and females respectively.

Within each study the difference between males and females seems to be significant, though the difference is not always in the same direction. In the Latin American countries, females seem to have greater returns on education than males. Because females in these countries tend to have lower education than males, they may have more scope to benefit productively from increases in education. Schultz's study also estimates higher returns to schooling for females than males, but the female returns to schooling that he found are much higher. Perhaps this has to do with the cross-country setting of his study.

McMahon's study yielded the opposite result: females had lower returns on education than males. This could be due to the fact that in the United States the gap between female and male schooling is small, and the methodology used in the study is somewhat different.

The foregoing review indicates, as do the summarizing tables in the extensive compilation by Psacharopoulos (1993, pp. 14-15, 44-47), that although female schooling appears to have a higher return than male schooling, the estimated returns show a large dispersion, and a direct global perspective is lacking.

A cross-country study that includes a larger number of countries at various stages of development and that estimates separate returns to schooling for males and females should be useful in shedding additional light on this important aspect. Such a study would be similar to that of Schultz (1994) in including schooling for both males and females in the model, and to that of Ram (1996) in using multiyear data and country-specific intercept dummies.

CHAPTER III

RESEARCH DESIGN

Although there are numerous ways to estimate returns to schooling, Mincer's schooling model provides a neat and simple framework that produces good results. Willis (1986, p. 526), who conducted an extensive survey on the theory underlying Mincer's schooling model and on the empirical research conducted with it, makes the following observation: ''as an empirical tool, the Mincer earnings function has been one of the great success stories of modern labor economics. It has been used in hundreds of studies using data from virtually every historical period and country for which suitable data exists. The results from these studies reveal important empirical regularities in educational wage differentials and the life cycle pattern of earnings.''

Because of its wide use, simplicity and theoretical appeal, Mincer's model was chosen as the main basis for the empirical analysis in this dissertation.

Mincer's Schooling Model

Jacob Mincer (1974, pp. 8-11) explained his model at some length. He reasoned that education is costly, because it is time consuming, and thus reduces the span of one's working life. It is also costly because income is deferred and because direct outlays are necessary to acquire education. Since education is costly, people will not invest in it unless, in addition to a possible direct consumption value, it results in a higher future income stream. This increase in future income must be enough to compensate, in present value, for the loss of income during the schooling period. For the two present values to be the same, a positive discount rate is necessary. Consistent with the view in the literature, Mincer used this discount rate as the internal rate of return on educational investment.

Mincer generated different versions of his schooling model by considering different sets of assumptions. For the model that will be used in this study, he offered the following assumptions.

First, a positive real discount rate prevails. This means that the nominal rate of return is greater than the inflation rate. Both theoretically and empirically this is a reasonable assumption.

Second, the marginal rate of return for the individual is constant at different levels of schooling and over time, which allows it to be treated as a parameter for the individual. What this means is that the return to every additional year of schooling will be the same as for the previous year, as well as for any other year of schooling. This implies that the average rate of return is also constant.

Third, investment costs will be considered as being only time costs. In other words, the only costs of education consist of the foregone earnings that the student could have earned had he or she chosen to work during those schooling years, rather than study. Schooling fees are ignored. This assumption greatly simplifies the model without affecting the spirit of the analysis. One way of interpreting this assumption is by considering that the direct outlays by students (such as tuition, books, etc.) are offset by their earnings in part-time jobs during schooling years.

Fourth, no further investment in human capital takes place after the schooling years. In other words, once the person gets a full-time job he or she will not make any further investments in learning. This implies that the individual does not incur any costs of on-the-job learning.

Fifth, the flow of earnings after the schooling years is assumed to be constant throughout the individual's working life, which greatly simplifies the model, without losing too much in terms of validity. Even though it is well known that the flow of income throughout the working life of an individual varies considerably, assuming it to be constant does not alter the analysis much, because what we are trying to account for here is the overall shift in the flow of income.

Sixth, the economy's productivity and earnings are assumed to be given over the worker's life-cycle. This is to focus on the effect that schooling has on productivity and earnings.

Seventh, no human-capital depreciation is assumed, which means that a fifty-year old worker and a forty-year old worker are equally productive if they have the same formal education. Thus, net investment in human capital should equal gross investment in human capital.

Finally, for this specific model it is assumed that one's working life is fixed, regardless of the years of schooling. Mincer based this assumption on the findings of a study by the U.S. Bureau of Labor Statistics (1970, Table E, p. A-11), which showed that working life tended to remain pretty constant between 45 and 47 years, even when the worker's education ranged anywhere from 8 years of schooling to more than 17 years.

Mathematically, Mincer (1974, pp. 8-11) constructed his model as follows:

Let

n = fixed span of earning life in years.

Y_s = annual earning of an individual with s years of schooling.

V_s = present value of an individual's lifetime earnings at start of schooling (with s years of schooling).

r = discount rate.

t = 0, 1, 2,..., n time, in years.

d = difference in the amount of schooling years, between two possible choices that are being considered.

e = base of natural logarithm.

Then, with continuous discounting, the present value of future earnings of a person who invests in s years of schooling will be

$$V_s = Y_s \int_s^{n+s} e^{-rt} dt = \frac{Y_s}{r} e^{-rs} (1 - e^{-rn}) . \tag{1}$$

Similarly, if the same person chooses to invest in a smaller number of years of schooling, say d less years of schooling, then that person would invest in s-d years of schooling, and the present value of future earnings would be

$$V_{s-d} = Y_{s-d} \int_{s-d}^{n+s-d} e^{-rt} dt = \frac{Y_{s-d}}{r} e^{-r(s-d)} (1 - e^{-rn}) . \tag{2}$$

Let $K_{s,s-d}$ be the ratio of annual earnings with s years to that with s-d years of schooling, and let r be interpreted as the internal rate of return.

Now, the person choosing between investing in s years of schooling and s-d years of schooling will only be willing to invest in s years of schooling, rather than s-d, if V_s is at least equal to V_{s-d}. Thus, this person would be indifferent between these two options if $V_s = V_{s-d}$. This indifference situation can be expressed by equating the present values from equations (1) and (2) and simplifying, which results in

$$K_{s,s-d} = \frac{Y_s}{Y_{s-d}} = \frac{e^{-r(s-d)}}{e^{-rs}} = e^{rd} . \tag{3}$$

Now suppose that the person in question is choosing between investing in s years of schooling and investing in no schooling at all. In such a case d=s, which enables us to substitute s for d in equation (3) and simplify, thus obtaining

$$K_{s,0} = \frac{Y_s}{Y_0} = e^{rs}. \tag{4}$$

Finally, in order to simplify the empirical study we can take logarithms on both sides of equation (4) and rearrange it as follows:

$$\ln Y_s = \ln Y_0 + rs. \tag{5}$$

Equation (5) shows the natural logarithm of earnings to be a strict linear function of years spent at school. Because it is a linear function, it can be econometrically studied through the usual regression methods. By applying this procedure to data for a cross section of m workers, the following econometric model emerges:

$$\ln Y_{si} = \ln Y_0 + rS_i + u_i \quad \text{for} \quad i = 1, 2, 3, \ldots, m \tag{6}$$

where Y_{si} denotes earnings of the ith worker, who has S_i years of schooling. Y_o denotes earnings with no schooling, which is the same for all workers; the coefficient r represents the average or common rate of return to schooling; and u_i is the error term, to account for random effects on income. Finally, m is the number of observations or workers in the sample.[1]

The key variables required for this model are schooling and income or output. How 'schooling' and 'output' are defined and what proxies are used to measure them empirically is open for discussion and are usually determined by what data are available. Other variables, such as physical capital stock and health-care spending, could be included in an extended model, but are not needed in the simplified Mincerian model. This model can also be augmented to distinguish between returns to formal education and on-the-job training.

Adapting Mincer's Model For This Study

First of all, Mincer constructed his model with an individual worker's education and earnings as the units of observation. Following Ram (1996), this model can be modified to place it in a cross-country context, where the units of observation are countries rather than workers, and extending it to include country-specific 'intercept' dummies. Thus, the coefficient for S_1 can be interpreted as representing the estimate for the average social rate of

[1] Mincer noted that this model has three distributional implications. First, if schooling were to have a symmetric, non-biased distribution, this would result in a positively skewed distribution of earnings. Mincer argued this is consistent with empirical evidence. Second, the larger the dispersion of schooling, the larger the dispersion and skewness of earnings. Finally, the higher the returns to schooling, the larger the earnings inequality and skewness.

return to schooling in the sample countries; and, in conjunction with the country-specific dummy variables, the constant term will now represent the estimate for $\ln(Y_o)$ for each country in the sample. The advantage of using country-specific ''intercept'' dummy variables is that this helps control for the many ''fixed'' cultural, social or other differences that may exist across countries.

In the simple Mincer model, S_i stands for years of schooling of the ith worker, who can be either a male or a female. But when this is applied in the cross-country context, S_i represents the average years of schooling of the entire labor force in a country at a point in time. Such an average is composed of the schooling of all males and all females in that population, which means that the average can be broken up into its two gender components. Thus, in principle it is possible to substitute S_i for its two gender components which are S_F for females and S_M for males, as Schultz (1994) did, to enter the schooling of females and of males separately. This modification enables us to obtain estimates of rates of returns to schooling by gender.

The result of these modifications to the basic Mincer schooling model provides us with the following regression model, which is used in this research:

$$\ln Y_{i,t} = a + r_F S_{F_{i,t}} + r_M S_{Mi,t} + \sum_{i=1}^{N-1} d_i C_i + u_{i,t} \qquad (7)$$

for $t=1,2,3,\ldots,T_i$ and $i=1,2,3,\ldots,N$

where:

$lnY_{i,t}$ = natural logarithm of the real gross domestic product per capita in country i during year t.

a = estimate for the natural logarithm of income in the base country with no education.

$S_{fi,t}$ = mean years of schooling per female in country i during year t.

$S_{mi,t}$ = mean years of schooling per male in country i during year t.

r_F = average social rate of return to the schooling for females.

r_M = average social rate of return to the schooling for males.

C_i = country specific dummy variable that will have value of 1 (one) for observations on country i and 0 (zero) otherwise.

N = number of countries in the sample

T_i = years of observation for the country i in the sample.

$u_{i,t}$ = error term for country i in year t.

Ram (1996) suggests that using the Mincer schooling model in a cross-country context reduces some of the more prominent problems of the model. First of all, the omission of an ''innate ability'' variable in the model is a serious one at the micro level. But when the model is applied on cross-country data, this ceases to be a problem, since it is unlikely that the average ``innate ability'' of workers in a country will differ much from that in other countries.

Second, on the micro level there may be a selectivity problem, in the sense that some workers may choose to study less due to good job opportunities that may pay them well. Other persons may choose to study further because of their ``comparative advantage'' in higher schooling. This could result in an underestimation of the returns to schooling. When using the cross-country approach, however, this problem is mitigated because of the fact that opportunities for such ``self-selection'' occur largely in intra-country choices.

Third, measurement of individual worker's post-school experience presents some problems at the micro level. With the cross-country data, though, this variable is not so relevant, since it is unlikely that variations in mean worker's income across countries can be systematically related to variations in on-the-job training and experience.

Apart from its global coverage, this model has the advantage that it provides separate estimates of returns to schooling for males and for females.

It may be noted that despite its advantages, the cross-country approach has some disadvantages also. First, it lacks a rigorous derivation of the sort presented above for the simple Mincer schooling model, especially when the schooling variable is separated into its male and female components. A second problem is that the opportunity-cost of schooling is not so clearly represented in this cross-country context. This is the main reason for the use of country specific ''intercept'' dummy variables, which help mitigate this problem, but do not eliminate it completely.

Whether applied on a cross-country basis or in large intra-country datasets involving individuals, the model can be criticized because of its single-equation format. This problem, however, can be mitigated by using the accumulated stock of human capital existing in a given year as the proxy for schooling, and using the current flow of income in that same year. As Ram (1996) argues, the flow of income in a given year is clearly dependent on the existing stock of human capital, but the flow of income in one year is likely to have only a small effect on the total stock of human capital that has been accumulated over many decades.

Finally, the fact that this model incorporates separate returns to schooling for males and females, but is based on the averaged income of both, is problematic. As long as the opportunity-cost of schooling Y_o is the same for males and females, the model should work fine. But if males are sacrificing more earnings than females during the schooling years, this model will overstate the rate of return to schooling for females and understate it for males, and vice-versa. By analyzing the results of the studies dealt with in the literature review, it does not become evident that the opportunity-costs for schooling of the genders are different. However, it is difficult to be sure that they are the same.

For this reason, an alternative form of this model is also used in this research, in which separate regressions are performed for males and for females. This model is as follows:

$$\ln Y_{i,t} = a_1 + r'_F S_{F i,t} + \sum_{i=1}^{N-1} d_{1i} C_i + u_{1i,t} \tag{8}$$

for t=1,2,3,...,T_i and i=1,2,3,...,N

$$\ln Y_{i,t} = a_2 + r'_M S_{M i,t} + \sum_{i=1}^{N-1} d_{2i} C_i + u_{2i,t} \tag{9}$$

for t=1,2,3,...,T_i and i=1,2,3,...,N

Again these two equations are variants of the cross-country Mincer schooling model. The main difference being that the schooling variable is either for one gender or the other. Although this approach eliminates the opportunity-cost problem mentioned above and resolves the effect of high collinearity between S_F and S_M in equation (7), it comes with its own disadvantages.

One important thing to note is that the income variable is the same in both equations (8) and (9). Ideally the income variable should have values specific for females in the female equation and values specific for males in the male equation. However, due to the present unavailability of data on cross-country gender-specific mean income, the same income variable had to be used in both cases. Since the income variable used is the result of the schooling of both genders, but only one gender is included in each equation, the rate of return to schooling may be overstated. In spite of this disadvantage, this model can still be used to complement equation (7) to study the differential in the returns to schooling across genders.

These models can be used to generate estimates by using the whole sample or by using subgroups of the sample, such as low-income countries, middle-income countries and high-income countries, in order to elucidate any patterns in the rate of return to schooling as countries become richer.

Due to diminishing returns to schooling, if there is a great discrepancy in schooling between males and females, there may also be a marked difference in the rate of return to schooling: the gender with the higher level of schooling would tend to experience a smaller rate of return to schooling. High-income economies tend to have a smaller difference in the levels of schooling across the genders than low-income economies. Thus, it may be expected that the higher the level of income in an economy, the smaller the differences in the rates of return to schooling.

CHAPTER IV

DATA AND METHODOLOGY

Sample Data and Sources

The target population is the world itself and the sample includes data from a set of countries that reasonably represents the whole world. Each observation consists of mean years of schooling of a country's female and male population and output per capita, in a given year.

The sample consists of 28 years of observations, from 1960 through 1987, on 82 countries. Table 5 on page 39 shows the list of countries that compose the sample. By using multiple-year data, the sample size is increased. More observations mean greater degrees of freedom, and this is important because of the large number of dummy variables used in this study to control for country-specific characteristics.

The proxies used for S_F and S_M are the average years of primary plus secondary education for females and males respectively, for the segment of the population of a country aged between 15 and 64 in a given year[2] . The data on

[2] ''Secondary'' schooling is divided into two cycles in many countries, such as junior high school and high school in the United States. When this occurs, both cycles are included in the secondary schooling variable.

schooling by gender is provided by Dubey and King (1994). This data consists of estimates of average primary and secondary years of schooling of males and females aged from 15 to 64. Dubey and King's data is an extension of the work done by Nehru, Swanson and Dubey (1993) on human capital stock across countries. Dubey and King (1994, pp. 4-5) refined this work by creating the gender distinction and by providing estimates of human capital for different age groups. Another improvement in Dubey and King's methodology was the relaxation of the assumption that dropout rates are the same at all levels of schooling. They account in their estimation procedure for variations in dropout rates across levels of schooling.

Dubey and King (1994, pp. 7-8) based their schooling estimates on data provided by UNESCO. UNESCO has been collecting all the necessary data from countries around the world on schooling and has attempted to make them comparable. To that effect, UNESCO created a standard classification for schooling figures presented by different countries, so that cross-country schooling comparisons became more meaningful.

Dubey and King's schooling estimates were based on UNESCO's data on school enrollments and on assumptions concerning the initial 'stock' of schooling. They took into account separate dropout rates and repeater rates for each year of schooling. Years of schooling due to repeating a grade were discounted, since they concluded that repeating a grade did not increase human capital by one year.

In dealing with the problem of 'depreciation' of human capital stock, Dubey and King considered that from the aggregate point of view of a country, human capital depreciates as people die, taking their knowledge and skills with them. So, like Nehru, Swanson and Dubey, they associated the probabilities of survival of persons of different age groups in order to come up with a human capital depreciation rate. Finally, Dubey and King were careful to take into account the changes in borders that occurred in various countries in the sample between the years when schooling took place and the year for which the human capital stock was being calculated.

Table 3 shows, according to Dubey and King's dataset, how the human capital stock has been changing over time. There are two relevant points that should be noted. First, across the span of the three decades both female and male schooling have risen significantly. Second, though average female schooling is shown to be smaller than for males, the difference has been decreasing over time. The table shows that the ratio of female to male schooling has been increasing at about 5 percentage points per decade from 67.1% to 77.6%. These trends should have an effect on the differential in the rates of return to schooling differential across the genders.

Table 3
Means and Standard Deviations of Female and Male Schooling Across Time.

	Female	Male	Female/Male	Observations
Full Sample:	3.86 (2.97)	5.33 (3.08)	72.4%	2279
1960-1969:	3.14 (2.96)	4.68 (3.29)	67.1%	818
1970-1979:	3.88 (2.91)	5.37 (2.99)	72.3%	820
1980-1987:	4.75 (2.82)	6.12 (2.71)	77.6%	641

Note: Standard deviations are in parenthesis below the means.

With respect to the differences across the three income differentiated subsamples, Table 4 shows, according to Dubey and King's dataset, how the human capital varies across groups of countries with different income levels. The countries in the sample are grouped into high-income, middle-income and low-income. Table 5 on page 39 shows the list of countries and how they are grouped. There are two relevant points that should be noted. First, both female and male schooling rise significantly as income rises. Second, though average female schooling is shown to be smaller than for males in every group, the difference is greatest in low-income countries and smallest in high-income countries.

Table 4 shows that the ratio of female to male schooling is 52.1% in low-income countries, while it is 85.2% in high-income countries. These differences should have an effect on the rate of return to schooling differential across the genders. Thus, we can expect to see different rate-of-return differentials for the three groups of countries.

Table 4
Means and Standard Deviations of Female and Male Schooling Across the Income Subsamples.

	Female	Male	Female/Male	Observations
High-Income:	7.34 (1.71)	8.61 (1.89)	85.2%	504
Middle-Income:	4.23 (2.64)	5.68 (2.75)	74.5%	805
Low-Income:	1.74 (1.60)	3.34 (2.12)	52.1%	970

Note: Standard deviations are in parenthesis below the means.

The proxy for income is the gross domestic product in constant-price ''international'' dollars per capita for each country and year. The data on income is based on an update (Penn World Table 5.5) that contains the estimates prepared by Summers and Heston (1991). These estimates, which have a good cross-country comparability, provide annual estimates of gross domestic product for a large number of countries, in constant 1985 international dollars, by using purchasing power parity exchange rates. This measure of income is presented in various forms, one of which is the ''real gross domestic product per capita''. This measure is fairly appropriate for this research.

Both the schooling and the income data have some limitations. First of all, both datasets rely on published figures of various governments throughout the world. In compiling such large datasets, big gaps in the data series inevitably emerge, and some extrapolation is perhaps unavoidable. Cross-country comparisons have other problems too. For an example on schooling, it is unlikely that 6 years of education in a Swiss school would yield the same human-capital stock as 6 years of education in Ethiopia. Furthermore, different governments use different ways to compute their educational figures, making cross-country comparisons difficult.

With regard to income, there are some problems too. Besides the inherent difficulty of adjusting for cross-country differences in relative price levels for major categories of goods and services, different countries assign different levels of resources to obtain estimates of their GDPs, which means that accuracy may vary greatly across nations. Furthermore, the varying sizes of underground economies across nations, may not be recorded at all in official figures. All these data limitations, which are common to most cross-country research, need to be kept in mind when interpreting the estimates.

Methodology

The ordinary least squares (OLS) method is used to estimate the modified Mincer schooling model that is specified in equation (7), namely:

$$\ln Y_{i,t} = a + r_F S_{F_{i,t}} + r_M S_{Mi,t} + \sum_{i=1}^{N-1} d_i C_i + u_{i,t}$$

for $t=1,2,3,\ldots,T_i$ and $i=1,2,3,\ldots,N$

The ''intercept'' country-specific dummy variables are meant to help control for differences in other ''fixed'' factors across countries. The coefficients r_F and r_M represent the average social returns to schooling for females and for males respectively in the sample countries. Thus, the coefficients r_F and r_M are the key in testing the null hypothesis that $r_F - r_M = 0$ or that r_F equals r_M. If the estimates support the hypothesis, this would suggest that there is no significant difference in the returns to schooling of females and males. A rejection of the hypothesis would suggest that there is a difference in returns to schooling across genders. In that case, if r_F is greater than r_M, the data would suggest that returns to schooling are greater for females than for males, and vice-versa. Regressions are estimated using the entire sample, including all the 82 countries and 28 annual observations for each country, from 1960 through 1987. Other regressions are also run for each of three subsamples, as listed in Table 5, consisting of low-income countries, middle-income countries, and high-income countries, in order to determine if differences exist across the three groups of countries in returns to schooling and the male-female differential. The first group consists of countries whose average real GDP per capita over the period 1960-1987 is lower than $2,000; the second group includes countries whose average real GDP per capita lies between $2,000 and $7,000; and the last group includes countries whose real GDP per capita is in excess of $7,000. The GDP figures used for this purpose are from an update (PWT 5.5) of Summers and Heston (1991).

Regressions are also performed separately for each decade in the sample. This should shed some light on trends in the rates of return to schooling and their differential across the genders.

Besides estimating equations (7), (8) and (9) for the full sample and for the three income differentiated subsamples, which directly provide estimates of the rates of return to schooling for males and females, the following reformulated version of equation (7) is also estimated:

$$\ln Y_{i,t} = a + r_1 S_{F_{i,t}} + r_2 S_{Ti,t} + \sum_{i=1}^{N-1} d_i C_i + u_{i,t} \qquad (10)$$

for $t=1,2,3,\ldots,T_i$ and $i=1,2,3,\ldots,N$

where $S_{Ti,t} = S_{Fi,t} + S_{Mi,t}$. This regression is estimated so as to facilitate a direct test of the null hypothesis of

Table 5
List of Countries Included in the Study and Their Grouping by Income Level.

High income group	Average GDP per capita	Middle income group	Average GDP per capita	Low income group	Average GDP per capita
United States	13736	Venezuela	5609	Paraguay	1771
Switzerland	12763	Singapore	5517	Thailand	1700
Canada	11756	Spain	5270	Philippines	1642
Australia	10224	Ireland	5098	Bolivia	1569
Sweden	9655	Uruguay	4703	Morocco	1568
Norway	9384	Cyprus	4493	Sri Lanka	1512
France	9116	Mexico	4400	Cote d'Ivoire	1492
Denmark	9049	Greece	4193	Mozambique	1402
Germany	8976	Argentina	4166	Angola	1227
Iceland	8915	Iraq	3809	Senegal	1194
Netherlands	8861	Chile	3725	Cameroon	1185
United Kingdom	8636	Syria	3596	Sierra Leone	1169
Belgium	8496	Portugal	3388	Liberia	1161
Finland	8051	Malaysia	3209	Zimbabwe	1158
Italy	7790	Brazil	3179	Honduras	1153
Austria	7701	Costa Rica	3168	Egypt	1132
Japan	7335	Iran	2986	Pakistan	1125
Israel	7102	Mauritius	2922	Indonesia	1122
		Peru	2879	Zambia	1088
		Panama	2723	China	1057
		Colombia	2710	Sudan	1029
		Turkey	2625	Ghana	1005
		Jamaica	2527	Nigeria	995
		Korea	2345	Haiti	899
		Ecuador	2274	Madagascar	892
		Tunisia	2252	Kenya	808
		Algeria	2247	Bangladesh	644
		Guatemala	2149	Rwanda	594
		Jordan	2037	Malawi	541
				Mali	467
				Burundi	436
				Zaire	414
				Tanzania	398
				Uganda	358
				Ethiopia	314

Note: Countries are grouped by their GDP per capita in constant 1985 international dollars, averaged for the period 1960-1987.

gender-specific rate-of-return equality, namely $r_F = r_M$. Notice that

$$r_1 S_{Fi,t} + r_2 S_{Ti,t} = r_1 S_{Fi,t} + r_2(S_{Fi,t} + S_{mi,t})$$

and so

$$r_1 S_{Fi,t} + r_2 S_{Ti,t} = (r_1 + r_2) S_{Fi,t} + r_2 S_{mi,t}.$$

Thus, if $r_1 = 0$ then $r_F = r_M = r_2$, which would imply that returns to schooling are the same for females as for males. Therefore, a statistically significant r_1 means that the null hypothesis, which states that returns to schooling are the same for females and for males, must be rejected, and its sign and magnitude indicate directly the excess of the rate of return to female schooling over that of male schooling.

Although OLS estimation of such models has well-known limitations, it does not seem likely that the broad character of the estimates is seriously affected by these. The aspect concerning the single-equation format was mentioned in Chapter III; that relating to the possibility of heteroscadasticity should be mitigated by the use of country-specific intercept dummies; and that pertaining to the possibility of autoregressive error structure for intra-country observations seems relatively minor.

CHAPTER V

THE MAIN RESULTS

Estimates of Returns to Schooling

Table 6 presents the estimates, for the entire sample as well as for the three income-based subsamples, of the adapted schooling model as shown in equation (7). Using the full multiyear-multicountry sample, female and male returns to schooling are estimated to be 26% and -6% respectively. Even though larger returns to schooling are expected for females than for males, returns to schooling are not expected to be so large for females and a negative rate of return for males appears implausible.

Comparing these estimates to those obtained by Schultz (1994), it is noteworthy that his estimates for female returns to schooling are also very large, though not nearly as large as the estimates in this research. For his 1970 regression, female return to schooling is estimated to be 16.9%, and for his 1980 regression it is 18.2%. He also found that male returns to schooling are lower than for females, though not as low as the ones in this study. For his 1970 regression male return to schooling is estimated to be 9.8%, and for his 1980 regression it is 11.0%.

Table 6
Estimates of the Adapted Schooling Model as Shown in Equation (7) From Multiyear Multicountry Data.

	Constant term	Coefficient of S_F	Coefficient of S_M	R^2 (F)	B (N)
Full Sample:	7.49 (111.82)	0.26 (27.39)	-0.06 (-6.61)	0.97 (732.03)	2279 (82)
High-income:	6.49 (37.83)	0.32 (27.48)	-0.02 (-1.66)	0.76 (79.14)	504 (18)
Middle-income:	7.88 (183.88)	0.19 (10.57)	0.03* (1.52)	0.80 (104.55)	805 (29)
Low-income:	6.72 (113.55)	0.21 (7.25)	-0.06 (-2.58)	0.86 (154.83)	970 (35)

Note: All estimates are significant at the 1% level, except those marked with *, which are not significant even at the 5% level. ``B'' denotes the number of observations.
The numbers in parenthesis under the estimates denote the corresponding t-value.
As equation (7) indicates, appropriate country-specific ``intercept'' dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_o$ are not a focus of the study.

The differences in the results of the two studies could be due to the fact that Schultz uses ``expected years of schooling'' enrollments lagged by a decade, while this study uses the actual estimated years of schooling. Another possible cause for this difference could have been the fact that Schultz did not use country-specific ``intercept'' dummy variables to control for fixed differences across countries. However, regressing equation (7) without the country-specific ''intercept'' dummy variables yields estimates for the rates of return to schooling for females and males of 34% and -8% respectively. Since these results fail to be any closer to those reported by Schultz, it can be concluded that the dummy variables are not the main cause of the discrepancies.

It can also be seen in Table 6 that in each of the subgroups the rate of return to schooling is greater for females than for males. As with the full sample, the estimate for female returns to schooling is higher than expected and positive, while that for males is lower than expected in all three cases and positive only for the middle-income countries and even that estimate is low and insignificant. We can tentatively infer from this that returns to schooling tend to be greater for females than for males, regardless of the level of income of a country.

Since the estimates can be interpreted as social rates of return, a larger-than-expected rate of return to schooling for females could be accounted for by the indirect effects of female education on the health and education of children.

However, it is possible that the real reason for these unexpected results is that the opportunity-cost of education varies across the two genders. Since equation (7) assumes that the opportunity-cost of education is the same for both sexes, this could cause the estimates for returns to schooling for females to be overestimated and that for males to be underestimated, if the male opportunity-cost for schooling is higher.

Another potential reason for these unexpected results might be that only a part of the adult schooling is used in the model. Dubey and King's data consists of years of primary and secondary schooling alone. Higher education is not included, and it might well be that including it would make a significant difference in the results.

Figure 1
Estimated Gender Differentials in Rates of Return to Schooling Across Income Subsamples, From Equation (7).

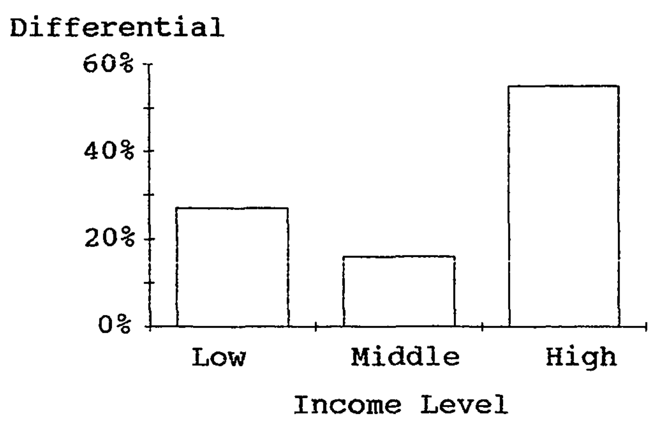


Figure 1 presents the difference in the estimates for the rates of return to schooling for females and males that are presented in Table 6. Although these differentials are probably overstated, an interesting pattern emerges from them. It suggests that as an economy increases its level of income and education, it may first witness a reduction in the gender differential in the rates of return to schooling; but as the level of income continues to increase, the differential will start rising again.

The first trend can be explained by the fact that as countries change from low-income to middle-income, schooling differentials across gender tend to become smaller, as Table 4 suggests. As a result of this, due to diminishing returns to schooling, the difference in the rate of return to schooling across the genders will tend to narrow down as well. The second trend, however, cannot be explained this way. The rising discrepancy could be the result of changing opportunity-costs as countries progress into high-income levels.

With regard to the changes across time in the rates of return to schooling and their difference across gender, Table 7 shows the results obtained by estimating equation (7) separately for each decade in the sample. There are a few things worth mentioning about the estimates shown in Table 7. To begin with, the opportunity-cost of schooling appears to rise over the three decades (at least in the base countries). That is to be expected, since real incomes have been rising as well.

The female rates of return to schooling are positive and decrease over time. That is to be expected as well, since female schooling has been rising faster than male schooling over time, as Table 3 shows. Male rates of return to schooling are all negative, with two out of the three estimates being small and insignificant. That is consistent with the findings shown in Table 6, and probably for the same reasons.

Table 7
Estimates Across Time of the Adapted Schooling Model as Shown in Equation (7) From Multiyear Multicountry Data.

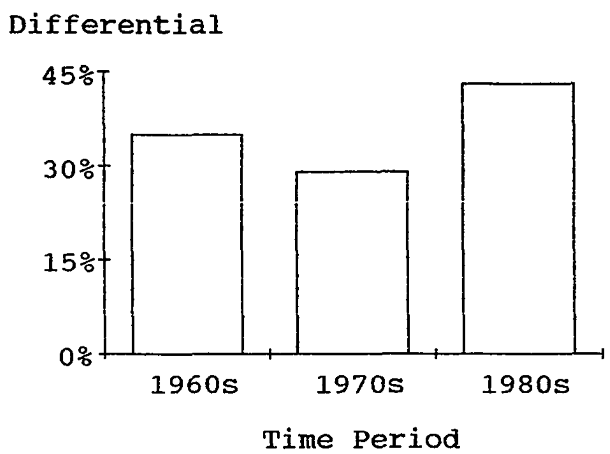
	Constant term	Coefficient of S_F	Coefficient of S_M	R^2 (F)	B (N)
1960-1969	6.20 (40.27)	0.33 (16.10)	-0.02* (-1.27)	0.99 (1025.41)	818 (82)
1970-1979	7.48 (119.78)	0.25 (11.14)	-0.04* (-1.69)	0.99 (844.67)	820 (82)
1980-1987	10.35 (42.01)	0.19 (4.01)	-0.24 (-4.76)	0.99 (485.53)	641 (82)

Note: All estimates are significant at the 1% level, except those marked with *, which are not significant even at the 5% level. ''B'' denotes the number of observations.
The numbers in parenthesis under the estimates denote the corresponding t-value.
As equation (7) indicates, appropriate country-specific ''intercept'' dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_0$ are not a focus of the study.

However, it is even more unlikely that the rate of return to schooling for males should be negative and as large as the estimate for the 1980-1987 period. A probable cause for these negative results, as with those in Table 6, is the possibility that the opportunity-cost for schooling is in fact different for the two genders. It could be that the difference in the opportunity-cost increased during the last decade, which would explain the large and negative estimate for males in that period.

Figure 2 demonstrates the pattern through time regarding the difference in the rates of return to schooling across the genders.

Figure 2
Estimated Gender Differentials in Rates of Return to Schooling Across Time Subsamples, From Equation (7).



The difference first falls from 35% to 29% and then rises again to 43%. Again, that finding is rather puzzling, since the difference in average schooling has decreased throughout the three time periods, as Table 3 shows. The smaller the difference in schooling, the smaller should be the difference in the rates of return to schooling. While this reasoning holds between the first two decades, it does not hold between the last two. Again, this could be due to the assumption of equal opportunity-costs for both genders.

Formal Test of the Hypothesis

Although Table 6 consistently indicates female rates of return to be substantially higher than those for males, it is of some interest to conduct a formal test of the difference. For that purpose, regressions are estimated using equation (10). Table 8 shows the estimates obtained with these regressions.

With the full sample, a value for r_1 of 0.33 is obtained with a t-statistic of 17.84. This value is significant at the 1% level. Thus, this test clearly rejects the null hypothesis that the rates of return for males and females are the same, and suggests the excess of the rate of return to female schooling over that to male schooling is of the order of 33%, which is what Table 6 also indicates without a formal test.

This conclusion is also supported by regressions for the three subgroups. In each case the null hypothesis (that returns to schooling for the two genders are the same) is rejected at a level of significance of 1%. We can therefore conclude that returns to schooling are different for females and males for economies at all stages of development. Furthermore, the excess of the female rate of return to schooling over the male rate of return appears in all cases.

Table 8
Estimates for Test of Gender Differential in Returns to Schooling Using Equation (10) From Multiyear Multicountry Data.

	Constant term	Coefficient of S_F	Coefficient of S_T	R^2 (F)	B (N)
Full Sample	7.49 (111.82)	0.33 (17.84)	-0.06 (-6.61)	0.97 (732.03)	2279 (82)
High-income	6.49 (37.83)	0.34 (16.66)	-0.02* (-1.66)	0.76 (79.14)	504 (18)
Middle-income	7.88 (183.88)	0.16 (4.44)	0.03* (1.52)	0.80 (104.55)	805 (29)
Low-income	6.72 (113.55)	0.27 (5.27)	-0.06 (-2.58)	0.86 (154.83)	970 (35)

Note: All estimates are significant at the 1% level, except those marked with *, which are not significant even at the 5% level. ``B'' denotes the number of observations.
The numbers in parenthesis under the estimates denote the corresponding t-value.
As equation (10) indicates, appropriate country-specific ``intercept'' dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_0$ are not a focus of the study.

<u>Other Regressions</u>

To compensate partially for the opportunity-cost problem inherent in equation (7), separate regressions are estimated for female and male schooling according to equations (8) and (9), which are:

$$\ln Y_{i,t} = a_1 + r'_F S_{Fi,t} + \sum_{i=1}^{N-1} d_{1i} C_i + u_{1i,t}$$

for t=1,2,3,...,T_i and i=1,2,3,...,N

$$\ln Y_{i,t} = a_2 + r'_M S_{Mi,t} + \sum_{i=1}^{N-1} d_{2i} C_i + u_{2i,t}$$

for t=1,2,3,...,T_i and i=1,2,3,...,N

As already explained, given the limitations imposed by the availability of data, the income variable used in both equations is the same. By running separate regressions, however, the opportunity-cost problem is mitigated, and the consequences of high collinearity between S_F and S_M in equation (7) are mitigated. Tables 9 and 10 show the estimates obtained with these regressions.

For the full sample, female and male returns to schooling are estimated to be 21% and 16% respectively. Although high, these results are closer to what may be expected. In fact, they are fairly similar to the estimates reported by Schultz (1994, p. 53). As for the subgroups, the regressions for high-income and for low-income countries produce estimates for returns to schooling that are greater for females than for males. With the high-income countries the difference is large, and female and male returns to schooling are estimated to be 31% and 8% respectively. For low-income countries the estimates are respectively 14% and 10%. Finally, the middle-income group returns to schooling are estimated to be the same for both sexes at 22%.

Table 9
Estimates of Female Returns to Schooling Based on Equation (8) From Multiyear Multicountry Data.

	Constant term	Coefficient of S_F	R^2 (F)	B (N)
Full Sample	7.35 (114.27)	0.21 (40.65)	0.96 (726.31)	2279 (82)
High-income Countries	6.28 (52.82)	0.31 (28.07)	0.76 (83.09)	504 (18)
Middle-income Countries	7.90 (189.68)	0.22 (32.67)	0.80 (107.89)	805 (29)
Low-income Countries	6.63 (138.38)	0.14 (13.83)	0.86 (158.11)	970 (35)

Note: All estimates are significant at the 1% level. The numbers in parenthesis under the estimates denote the corresponding t-value. ''B'' denotes the number of observations.
As equation (8) indicates, appropriate country-specific ''intercept'' dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_o$ are not a focus of the study.

Interestingly, the base-country opportunity-costs estimated by using equations (8) and (9), which are presented as the constant terms in Tables 9 and 10, suggest that the female opportunity cost may be lower in several cases. Again, this supports the argument presented above that opportunity cost differences may be the main cause of the high differential found in the rates of return to schooling across the genders in estimates of equation (7).

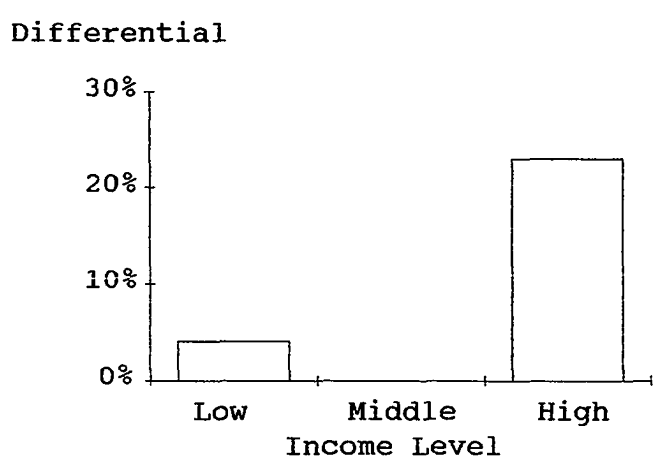
Table 10
Estimates of Male Returns to Schooling Based on Equation (9) From Multiyear Multicountry Data.

	Constant term	Coefficient of S_M	R^2 (F)	B (N)
Full Sample	7.78 (101.52)	0.16 (27.00)	0.95 (545.63)	2279 (82)
High-income Countries	8.58 (34.94)	0.08 (3.88)	0.38 (16.29)	504 (18)
Middle-income Countries	7.87 (171.76)	0.22 (28.98)	0.77 (91.27)	805 (29)
Low-income Countries	6.56 (116.71)	0.10 (11.79)	0.85 (149.50)	970 (35)

Note: All estimates are significant at the 1% level. The numbers in parenthesis under the estimates denote the corresponding t-value. "B" denotes the number of observations.
As equation (9) indicates, appropriate country-specific "intercept" dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_0$ are not a focus of the study.

Figure 3 shows the difference in the estimates for the rates of return to schooling for females and males that are presented in Tables 9 and 10. It shows a pattern similar to that of Figure 1, albeit with smaller differentials overall and more plausible estimates of the rate-of-return. Again, it suggests that as an economy increases its level of income it will first witness a reduction in the gender differential in the rates of return to schooling; but as the level of income continues to increase, the differential will start rising again.

Figure 3
Estimated Gender Differentials in Rates of Return to Schooling Across Income Subsamples, From Equations (8) and (9).



The foregoing discussion indicates that although the magnitude of the gender differentials varies, the broad pattern of the estimates of equation (7) are similar to those of equations (8) and (9). Both suggest significantly higher rates of return to schooling of females than of males in the whole sample as well as in most subgroups.

Furthermore, both indicate a similar pattern for the rate of return differential across the three income subgroups.

Given the fact that King's dataset is so recent and comprehensive, it is worthwhile to compare the results with studies that used other datasets for schooling. In order to make this comparison, a regression can be run that does not discriminate schooling by gender. For this purpose, the total schooling variable S_T is used that combines the overall schooling of both females and males. The regression model used is the simple Mincer schooling model with cross-country dummy variables, which is:

$$\ln Y_{i,t} = a_3 + r_T S_{T_{i,t}} + \sum_{i=1}^{N-1} d_{3i} C_i + u_{3i,t} \qquad (11)$$

for $t=1,2,3,\ldots,T_i$ and $i=1,2,3,\ldots,N$

This model is very similar to the one used by Ram (1996). Thus, the estimates obtained from equation (11) can be compared with those presented by Ram (1996). Table 11 presents the results produced by this regression; these are very good in the sense that they are well within the range found in other such studies. The results compare remarkably well with Ram (1996), even though he used a different source for data on years of schooling.

Ram estimates a rate of return of 13% with the full sample, which is close to the 10% presented here. The similarities extend into the three income differentiated subgroups, with the only marked discrepancy arising in the high-income countries subgroup. Generally, he had slightly higher estimates for the rate of return to schooling in the low-income and middle-income countries, but a much lower estimate for the high-income countries.

Table 11
Estimates of Overall Returns to Schooling Based on Equation (11) From Multiyear Multicountry Data.

	Constant term	Coefficient of S_T	R^2 (F)	B (N)
Full Sample	7.40 (103.57)	0.10 (35.10)	0.96 (644.05)	2279 (82)
High-income Countries	5.80 (27.80)	0.17 (18.13)	0.62 (43.38)	504 (18)
Middle-income Countries	7.86 (183.00)	0.11 (32.04)	0.80 (104.94)	805 (29)
Low-income Countries	6.57 (124.91)	0.06 (12.91)	0.85 (154.05)	970 (35)

Note: All estimates are significant at the 1% level. The numbers in parenthesis under the estimates denote the corresponding t-value. ''B'' denotes the number of observations.
As equation (11) indicates, appropriate country-specific ''intercept'' dummy variables are used in all cases, but their estimates are not reported since cross-country differences in $\ln Y_o$ are not a focus of the study.

Overall, the results strongly support the idea that returns to schooling are greater for females than for males. Almost all estimates point in that direction, whether using the adapted schooling model presented in equation (7) or separate regressions for males and females according to equations (8) and (9). Even with the regressions on the subgroups of countries classified according to income, the results suggest greater returns to schooling for females than for males in most cases. The test for equality of returns to schooling for males and females rejects the null hypothesis at the 1% level in the full sample and all of the three country subsamples.

With regard to the differences across the three subsamples, a pattern emerges suggesting that differences in returns to schooling will depend on the level of income per capita in the country. The regressions show the highest difference on returns to schooling across gender in the high-income subgroup and the smallest difference in the middle-income subgroup. This suggests that as income per capita increases, the difference in returns to schooling across genders will first diminish, as the country in question emerges out of poverty. But as income continues to grow and the country becomes richer, the differential may begin to rise again.

When comparing the results of this study with those compiled by Psacharopoulos (1993) and with those presented by Psacharopoulos & Ng (1992), Schultz (1994), McMahon (1991), and Hertz & Khandker (1991), some differences and similarities emerge. Like Psacharopoulos & Ng and Hertz & Khandker and many other studies, this work indicates higher returns to schooling for females than for males. However, gender differences indicated by this study are typically much larger. Even when separate regressions are performed for males and for females, the differences in returns to schooling are much greater than previous studies indicate. On the other hand, the total returns to schooling that are shown in Table 11, are very similar to those found in other studies that do not discriminate by gender, such as that of Ram (1996), which suggests that the data sources used here are reasonable.

As already stated, since cross-country data on income by gender is simply not yet available, the study had to use average income (GDP per capita). The weakness of that procedure is evident, and it is possible that the results of this study tend to overstate the gender difference in returns to schooling.

A good follow-up on this study, when the availability of data permits, would be one in which there is an income proxy specific for each gender. Such a procedure would probably generate more reliable estimates.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary of Research

The importance of education for economic growth has been stressed by economists for a long time. Consequently, a large amount of research has been done on this relationship, both at the theoretical and the empirical levels. All the same, large gaps still remain in our knowledge of this topic.

Given the low level of female schooling in many countries and the increasing participation of females in the labor force, comparisons of private and social returns to schooling of females and males assumes greater importance. Moreover, a global perspective on such comparisons seems particularly useful.

Thus, there is a need for studies that estimate returns to schooling by gender and that are based on an international sample. Although there have been many studies on gender-specific rates of return to schooling, just one of these, by Schultz (1994), seems to have been conducted on a cross-country basis, and even that used a somewhat special proxy for the level of schooling.

The purpose of this research is to use multicountry data to estimate the impact that female education has on aggregate output or income, and to find out how different, if at all, it is from the impact of male education.

Two models are used in this study, and both are based on the ''schooling model'' developed by Mincer (1974). Both are modified, as Ram (1996) did, by making countries, rather than individuals, as the units of observation. The basic model is modified along the lines suggested by Schultz (1994) to make it capable of generating separate estimates of the rates of return to schooling for females and for males. Estimates are also obtained by running separate regressions for females and males.

The models are estimated for the whole sample as well as three subsamples, where countries are classified as low-income, middle-income and high-income. The subsamples provide insights on diminishing returns to schooling and its effect on gender differences in the rate of return to schooling. The model is also estimated separately for each decade of observations in the sample, to throw some light on the changes over time of the rates of return to schooling.

The sample consisted of 28 years of data for 82 countries. The proxy for the schooling variable is the average number of years of primary and secondary schooling for females and males aged between 15 and 64. The data on schooling is provided by Dubey and King (1994).

The proxy for income is the gross domestic product in constant-price ''international'' dollars per capita for each country and year. The data on income is based on an update (Penn World Table 5.5) of the estimates prepared by Summers and Heston (1991).

The results obtained with both models clearly indicate that the rate of return to schooling is larger for females than for males. This is true for regressions based on the whole sample as well as for most of the subgroups. The results based on the first model indicate that female returns to schooling are substantially larger than those reported in previous studies. On the other hand, male returns to schooling are shown to be low or negative, which appears somewhat implausible.

There are a few possible explanations for these results. One has to do with the possibility that opportunity-costs for female and male schooling are not the same. If they in fact are not the same, the model will tend to overstate the rate of return to schooling for the gender with the lowest opportunity-cost, and vice-versa. The second model, which does not assume equal opportunity-costs for males and females, however, has the drawback that the income variable is the average for both genders. In spite of that, the second model produced more plausible results, although the estimates tend to be a bit high.

A second possible cause for these results could be the fact that the proxy for schooling used does not include higher education.

Substantial collinearity between male and female schooling could also be a cause of the patterns observed in the estimates of equation (7). That problem is also resolved through the estimates of equations (8) and (9), which yield much more plausible estimates than (7).

Regarding the patterns in the income-based subsamples, it appears that the excess of the female rate of return over the male rate tends to decrease as countries progress from low-income to middle-income status. This may be because as countries develop, schooling differentials across gender tend to narrow down. Due to diminishing returns to schooling, the difference in the rate of return to schooling across the genders will tend to become smaller.

However, the excess of the female rate of return over the male tends to rise again as countries further progress from middle-income to high-income levels. The rising differential could be the result of changing opportunity-costs as countries progress into high-income levels.

Concerning the patterns across the three decades, the excess of the female rate of return over the male rate tends to decrease from the 1960s to the 1970s, but then increased from the 1970s to the 1980s. A possible explanation for this could be the same as the one mentioned for the income-based subsamples. As the world economy grew from the 1960s to the 1970s, gender-specific education differentials may have narrowed, resulting in more similar rates of return across gender. Further progress into the 1980s could have resulted in widening opportunity-cost differentials, resulting in an overestimation of female returns to schooling and an underestimation of male returns to schooling.

Suggestions for Future Research

This research has a few significant limitations that affect the results. One primary limitation is the one imposed by having the same income variable for males and females. When an international dataset on income per adult that discriminates by gender becomes available, gender-specific income measures should be used along with gender-specific schooling variables. In that case, separate regressions could be run for males and for females, but each with its own income variable, thus reducing the omitted-variable problem in the second model used in this research. By not having to use a common income for both genders, as the first model does, the need to assume equal opportunity-costs for schooling would be eliminated. Thus, such a gender-specific income data would eliminate a major limitation of this research.

Another suggestion is that similar exercises be conducted on the basis of gender-specific schooling data that includes higher education as well as primary and secondary schooling. Again this will have to wait until such a database becomes available.

WORKS CONSULTED

Becker, G. S. *Human Capital.* New York: National Bureau of Economic Research; Columbia University Press, 1964.

Denison, E. N. ``The Sources of Economic Growth in the U.S. and the Alternatives Before Us.'' Supplementary Paper No. 13, New York: Committee for Economic Development, 1962.

Dubey, A. and King, E. ``A New Cross-Country Education Stock Series Differentiated by Age and Sex.'' Working Paper, Washington DC: World Bank, August 1994.

Fabricant, S. ``Basic Facts on Productivity.'' Occasional Paper No. 63, New York: National Bureau of Economic Research, 1959.

Griliches, Z. ``Estimating Returns to Schooling: Some Econometric Problems.'' *Econometrica*, 45 (January 1977): 1-22.

Hertz, B. K., and Khandker, S. R. ``Women's Work, education and Family Welfare in Peru.'' Discussion Paper 116, Washington DC: The World Bank, 1991. (ERIC Document Reproduction Service No. ED 338 525).

Lucas, R. E. ``On the Mechanics of Economic Development.'' *Journal of Monetary Economics*, 22 (July 1988): 3-42.

Mankiw, N. G., Romer, D., & Weil, D. N. ``A Contribution to the Empirics of Economic Growth.'' *Quarterly Journal of Economics*, 107 (May 1992): 407-437.

McMahon, W. W. ``Relative Returns to Human and Physical Capital in the U.S. and Efficient Investment Strategies.'' *Economics of Education Review*, 10 (December 1991): 283-296.

Mincer, J. *Schooling, Experience, and Earnings.* New York: National Bureau of Economic Research; Columbia University Press, 1974.

Nehru, V., Swanson, E., and Dubey, A. ``A New Database on Human Capital Stock: Sources, Methodology, and Results.'' Working Paper No. 1124, Washington, DC: The World Bank, 1993.

Psacharopoulos, G. *Returns to Education, an International Comparison*. Washington DC: Jossey-Bass, 1973.

________. ``Returns to Education: A Further International Update and Implications.'' *The Journal of Human Resources*, 20 (Fall 1985): 583-604.

________. ``Returns to Investment in Education: a Global Update.'' Working Paper No. 1067, Washington DC: The World Bank, January 1993.

Psacharopoulos, G., and Arriagada, A. M. ``The Educational Composition of the Labor Force: An International Comparison.'' *International Labour Review*, 125 (September-October 1986): 561-574.

________. ``The Educational Composition of the Labour Force: An International Update.'' *Journal of Educational Planning and Administration*, 6 (April 1992): 141-159.

Psacharopoulos, G., and Ng, Y. C. ``Earnings and Education in Latin America. Assessing Priorities for Schooling Investments.'' Working Paper 1056, Washington, DC: The World Bank, December 1992.

Psacharopoulos, G., and Tzannatos, Z. *Case Studies on Women's Employment and Pay in Latin America.* Washington DC: The World Bank, 1992.

Psacharopoulos, G., and Woodhall, M. *Education for Development*, New York: Oxford University, 1985.

Ram, R. ``Economic Development and Rates of Return to Schooling: Some Estimates From Multicountry Data.'' Forthcoming in *Economic Development and Cultural Change* (1996).

Schultz, T. P. ``The Role of Education and Human Capital in Economic Development: an Empirical Assessment.'' Center Discussion Paper No. 670, New Haven, Connecticut: Economic Growth Center, Yale University, August 1992.

________. ``Human Capital Investment in Women and Men: Micro and Macro Evidence of Economic Returns.'' Occasional Paper No. 44, San Francisco, CA: International Center for Economic Growth, 1994.

Schultz, T.W. *The Economic Value of Education*. New York: Columbia University Press, 1963.

________. *Investment in Human Capital*. New York: Free Press, Macmillan, 1971.

Solow, R. M. ''A Contribution to the Theory of Economic Growth.'' *Quarterly Journal of Economics*, 70 (February 1956): 65-94.

________. ''Technical Change and the Aggregate Production Function.'' *Review of Economics and Statistics*, 39 (August 1957): 312-20

Summers, R. and Heston, A. ''The Penn World Table (Mark 5): An Expanded set of International Comparisons, 1950-1988.'' *Quarterly Journal of Economics*, 106 (May 1991): 327-368.

U.S. Department of Labor, Bureau of Labor Statistics. ''Educational Attainment of Workers.'' Special Labor Force Report No. 83, March 1970.

Willis, R. J. ''Wage Determinants: a Survey and Reinterpretation of Human Capital Earnings.'' In: Aschenfelter O., Layard R. (eds). *Handbook of Labor Economics*. North-Holland, Amsterdam, 1986: 525-601.