EDUCATION AND GROWTH: IS IT QUALITY OR QUANTITY THAT MATTERS?

by

Dongxia Xu
B.A., University of International Business and Economics, 2002

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
Department
of
Economics

SIMON FRASER UNIVERSITY

SPRING 2006

APPROVAL

Name: **Dongxia Xu**

Degree: **M. A. (Economics)**

Title of Project: **Education And Growth: Is It Quality Or Quantity That Matters?**

Examining Committee:

Chair: **Clyde Reed**

Brian Krauth
Senior Supervisor

Jane Friesen
Supervisor

Simon Woodcock
Internal Examiner

Date Approved: Friday, April 7, 2006

Abstract

The literature on growth effects of education tend to focus on quantity measures of schooling, leading to some findings that cast doubt on the role of human capital in economic growth. This paper, following Hanushek and Kimko (2000), investigates the importance of schooling quality constructed from student performance on international tests in determining cross-country differences in growth rates. The results show that schooling quality is positively and significantly related to growth rates. This relationship is robust to alterations in the controlling variables and not driven by some East Asian countries. Comparison of alternative measures of schooling quality provides further support that quality measures based on cognitive skills are better measures of schooling quality than school inputs. In contrast, schooling quantity measures lose significance when quality measures are included. Finally, causality tests indicate the strong relationship between schooling quality and growth might be subject to reverse causality.

Keywords:

Education, Growth, Quality of schooling, Quantity of schooling

Acknowledgements

First, I am greatly indebted to my senior supervisor Professor Brian Krauth for giving me continuous support and valuable advice. He was always there to talk about my ideas and to help me through the problems and difficulties I encountered in the process of writing this paper. I also would like to thank my supervisor Professor Jane Friesen for her helpful suggestions and comments.

A special thanks goes to my parents for their unconditional support and encouragement in my academic pursuit. I also thank my husband for being a patient listener when I was frustrated and for showing confidence on me.

Last, I am grateful to my friends Yi Xue for his suggestions and Patrisha de Leon-Manlagnit for providing helps in searching data.

List of Contents

List of Tables

1. Introduction

While the role of human capital in economic growth has been emphasized by endogenous growth theory, the empirical literature using quantity measures of schooling as a proxy for human capital shows little consensus on the impacts of human capital on economic growth. Hanushek and Kimko (2000) point out the importance of quality measures of schooling constructed from international test scores in determining cross-country differences in GDP growth rates. This paper, following Hanushek and Kimko (2000), investigates the growth effect of schooling quality based on updated data.

First, I construct schooling quality indexes based on student performance on international tests at various years. While Hanushek and Kimko's quality measures consider test years up to 1991, I extend the test years to 1995. Second, I estimate cross-sectional growth regressions. These regressions show that schooling quality indexes are positively and significantly associated with economic growth and schooling quantity measures lose significance when schooling quality is controlled for. These results are consistent with the findings in Hanushek and Kimko (2000).The estimated relationship between quality measures and growth is found to be robust to alterations in the controlling variables and to the exclusion of East Asian countries. The comparisons between the constructed quality measures and school inputs---alternative quality measures--- indicate that the coefficients on the constructed quality indexes retain the positive sign and the significance with the inclusion of schooling inputs. On the other hand, school input variables either have the unexpected signs or become insignificant when direct quality measures are controlled for. This finding provides evidence that quality measures based on cognitive skills are better measures for human-capital quality. Finally, one may be concerned about reverse causality, i.e., if higher growth rates lead to better schooling quality through the channel of government investment in schooling resources. To investigate this issue, I estimate the determinants of schooling quality with the same estimation technique employed in Hanushek and Kimko (2000) by relating test scores at year t to the relevant school resources at year $t-1$. In contrast to the results in Hanushek and Kimko (2000) which finds that increases in school resources can not improve test scores, my results based on updated school resource data show some evidence that school resources could affect test performance. Thus, reverse causality is a concern for the coefficient estimates on quality measures in the cross-sectional regressions.

1.1 Literature Review

A great number of empirical studies have been made to examine the growth effect of human capital.

Among the earlier work, Barro (1991) in his study of economic growth for 98 countries finds that the initial secondary enrollment rates are positively related to the growth rates in GDP per capita. He also includes the student-teacher ratio as a control for schooling quality and finds that the student-teacher ratio in primary schools is negatively related to growth while the student-teacher ratio in secondary schools has an insignificant positive association with growth. In addition, Mankiw, Romer and Weil (1992) estimate the augmented Solow model for three samples of countries[1] and they report a positive and significant relationship between secondary school enrollments and growth rates in all three samples.

In contrast, Benhabib and Spiegel (1994) find human capital accumulation proxied by the average years of schooling enters insignificantly with a negative sign in cross-country growth regressions. Pritchett (1996) provides further supports to the findings in Benhabib and Spiegel (1994) by demonstrating that the estimated negative relationship between educational attainments and economic growth is robust to choice of sample and estimation technique. While previous findings are based on cross-country data, an insignificant growth effect of human capital, proxied by the average years of schooling, is also found by Islam (1995) through a panel-data approach.

However, a later study by Krueger and Lindahl (2001) reaffirms the positive relationship between educational attainments and growth after considering the measurement errors in educational variables in Benhabib and Spiegel's (1994) paper.

Although the findings on the growth effect of human capital are controversial, a common feature of these studies is that they, at least to some extent, ignore the quality of education while using only some quantity measures of education such as school enrollment rates or average years of schooling to proxy human capital. Hanushek and Kimko (1995 and 2000) emphasize the important role of labor-force quality in determining growth rates. They find the direct quality measures of schooling constructed from student performance on international examinations have a strong causal effect on

[1] These three samples of countries are the 98 non oil-exporting nations, the 75 non oil-exporting nations with more reliable data and the 22 OECD countries.

GDP growth rates. However, quantity measures of schooling enter insignificantly whenever labor-force quality is considered in the regressions.

This paper follows Hanushek and Kimko (2000) estimating how quality measures based on expanded set of test scores affect economic growth rates in a set of cross-sectional countries, compared to the effect of quantity measures. The remainder of the paper is organized as follows. Section 2 describes theory, model, data sources and the construction of schooling quality measures. Section 3 estimates the effects of quantity and quality of schooling on growth. Section 4 investigates the robustness of quality effect on growth. Section 5 compares the alternative measures of schooling quality. Section 6 examines the issue of reverse causality. Section 7 concludes.

2. Methodology

2.1 Theory and Model

In the Solow Growth Model or neoclassical growth theory, technological progress is taken as exogenous and human capital plays no role in the output production. Nevertheless, endogenous growth theory rejects the assumption of exogenous technological change and views human capital as an important factor in determining long-run economic growth.

Various models have been established to support endogenous growth theory. In the simple AK model, non-diminishing marginal return to aggregate capital that includes human capital leads to endogenous growth in output per capita (Rebelo, 1991). In the learning-by-doing model, endogenous growth comes from knowledge accumulation in the production process (e.g., Romer, 1986). In the technology adoption models, the stock of human capital may influence productivity through affecting a country's ability to innovate and develop new domestic technologies or to imitate and absorb advanced technologies from abroad (e.g., Nelson and Phelps, 1966; Romer, 1990a). These endogenous growth models provide theoretical foundation to the empirical work in attempting to examine the role of human capital in economic growth.

However, as mentioned in the previous section, the empirical literature using quantity measures of schooling as a proxy for human capital produces inconsistent results on the growth effect of human capital. Hanushek and Kimko (2000) argue that the results from the past studies which have paid more attention to quantity measures of education might be misleading if quality of education is a more important dimension of human capital than quantity. Thus, this paper following Hanushek and Kimko (2000) includes quality measures of schooling in the model specifications. The model specification used in this paper is as follows[2]:

$$\Delta Y_i = \beta_o + \beta_1 Y_{i,t-1} + \beta_2 S_i + \beta_3 QL_i + \beta_3 Z_i + \varepsilon_i$$

where ΔY_i is country *i*'s annual average growth rate from year *t-1* to year *t*; $Y_{i,t-1}$ is the initial GDP per capita of country *i*; S_i represents a quantity measure of education in country *i* and QL_i is the quality measure for country *i*. Z_i stands for other variables that may affect economic growth rate such as physical capital, trade and other government policies.

2.2 Data on Test Scores

Following Hanushek and Kimko (2000), I develop quality measures based on international tests in mathematics and science conducted over the past four decades, but I update the test year to 1995 while they include tests through 1991. Using tests in mathematics and science instead of tests on other subjects comes down to the important role of mathematics and science in guiding research and development activities and its role in fostering future engineers and scientists (Hanushek and Kimko, 2000).

A total of seven international tests[3] are considered in my quality measures. This sub-section gives a brief introduction to the international tests involved in this paper before going into details of the construction of quality measures.

[2] This model is a simple transformation of the "typical estimating equation" employed in Krueger and Lindahl (2001).
[3] Six tests are involved in Hanushek and Kimko's (2000) measures.

Seven international tests in mathematics and science were conducted on samples of primary and secondary students from various countries in the same age group[4] by the International Association for the Evaluation of Educational Achievement (IEA) and the International Assessment of Educational Progress (IAEP). Table 1 describes details for each study.

The IEA[5], established in 1958, organized five out of those seven international examinations. They are, in year order, the First International Mathematics Study (FIMS), the Six Subject Study[6] (the First International Science Study (FISS) is part of this research project), the Second International Mathematics Study (SIMS), the Second International Science Study (SISS), and the Third International Mathematics and Science Study (TIMSS)[7].

The International Assessment of Educational Progress (IAEP)[8], initiated in 1988, capitalizes on the materials and procedures of a major testing project in the United States, named the National Assessment of Educational Progress (NAEP). IAEP administered two international tests. The first IAEP study, an experimental test conducted in the U.S and five other countries in 1988, assessed 13-year-olds students' achievement in mathematics and science. The second IAEP study was conducted in the U.S and nineteen other countries in 1991, surveying students of 9- and 13-year-olds in the subjects of mathematics and science.

---

[4] Target population includes students of 9 or 10-year-olds, 13 or 14-year-olds and students at the final year of secondary school.

[5] More information on the history of IEA can be retrieved at http://www.iea.nl/brief_history_of_iea.html

[6] The six subjects are science, reading, literature, English and French as foreign languages, and civic education.

[7] TIMSS is not included in Hanushek and Kimko (2000).

[8] More information on the IAEP studies can be retrieved at http://fermat.nap.edu/html/icse/study_h.html

Table 1 International tests in mathematics and science

Tests	Years of data collection	Sponsor	Subjects	No. of Countries	Ages of Pupils
FIMS	1964-1966	IEA	Math	11	13, Final sec[a].
FISS[b]	1970-1971	IEA	Science	17	10, 14, Final sec.
SIMS	1980-1981	IEA	Math	20	13, Final sec.
SISS	1983-1985	IEA	Science	24	10, 14, Final sec.
First IAEP Study	1988	IAEP	Math Science	6 6	13 13
Second IAEP Study	1990-1991	IAEP	Math Science	20 20	9, 13 9,13
TIMSS	1995	IEA	Math Science	41 41	9, 13. Final sec. 9, 13. Final sec.

Notes: [a]Final sec. denotes the final year of secondary education.
[b]FISS is part of Six Subject Studies.
Source: Adapted from "Schooling quality in a cross-section of countries" (2001)

The main source of test scores used in this paper is Barro and Lee schooling quality dataset[9]. Barro and Lee (1996) have compiled cross-country data on average test scores from international comparative studies for students at different age groups. As the original test results are reported in the format of either number of items correct, or percent correct, or scores in proficiency scale with a mean 500 and a standard deviation of 100 over the range of 1 to 1000, Barro and Lee transformed all test scores into the percent-correct format for comparability between different tests (Barro and Lee, 2001). However, the average test scores for students of 9-year-olds and students at final year of secondary education[10] in the Third International Mathematics and Science Study (TIMSS) are not available in Barro and Lee's schooling quality data. I add them into the dataset separately[11]. With the expanded dataset, underlying test scores are available in 52

[9] Data for test scores are available from the World bank website: http://econ.worldbank.org
[10] Only test scores of 13-year-old students are available for TIMSS.
[11] See Appendix A for data sources.

countries[12], among which only the United States and United Kingdom participated in all tests.

2.3 Construction of Schooling Quality Measures

I construct a single index of schooling quality for each country, using the approaches employed in Hanushek and Kimko (2000). To the extent that the stock of human capital attended schools at different time spans, an index with the mixture of tests conducted at varying times could be a proper approximation to the aggregate level of human capital; however the underlying assumption is the quality of education systems evolves slowly over time, which is justified in some sense by the fact that teaching technology is comparatively stationary over a certain period and the turnover of teachers and other school personnel is slow (Hanushek and Kimko, 2000).

Following Hanushek and Kim (1995), quality measures are developed by constructing a weighted average of the normalized test scores. Test scores are normalized in two ways. The first normalization method is to convert every performance series (reflecting different age and test year) that has different group mean percent correct to a mean of 50. This is done by multiplying the original group mean of each test series τ, say m_τ, by $50/m_\tau$. This approach assumes a constant intertemporal mean in the international cognitive skill level measured by test scores in science and mathematics and that the participating countries are drawn randomly from the world distribution (Hanushek and Kim, 1995). The second normalization method draws on the fact that the United States participated in all international tests and U.S. National Assessment of Educational

[12] Only 39 observations are available in Hanushek and Kimko (2000), because TIMSS with most countries participated is not included in their data.

Progress (NAEP) has conducted mathematics and science tests to samples of students aged 9, 13 and 17 on a consistent basis at different times since its establishment in 1969[13]. This allows us to transform test scores of U.S students in international studies to mimic their scores in mathematics and science in the nearest year and age-group in NAEP[14]. Then performance of other countries is adjusted proportionately in accordance to relevant transformed U.S. scores. Denote the normalized average test score for country i in test τ as $S_{\tau i}$ .Then a weighted average is taken over all transformed test scores $S_{\tau i}$ with the weight for country i in test τ as $WT_{\tau i}$, where

$$WT_{\tau i} = \frac{1/\sigma_{\tau i}}{\sum_{\tau} 1/\sigma_{\tau i}}$$

($\sigma_{\tau i}$ is the normalized standard error[15] of mean test score for country i in test τ.).

This implies that a smaller weight is assigned to the average test score that has a larger standard error. Hanushek and Kim (1995, p. 19) argue that "a high standard error conveys less accurate information about the country's performance." So, the quality index for country $i = \sum_{\tau} S_{\tau i} * WT_{\tau i}$.

Quality Index 1 (QL1) and Quality Index 2 (QL2) are constructed using the first and second normalizing method respectively. Twenty-two test performance series that are available for various subsets of participating countries are combined into a single

---

[13] NAEP test scores for various years are published in *Digest of Education Statistics (1997)* by the U.S. Department of Education.

[14] As FIMS was conducted by IEA in 1964 predated the establishment of NAEP, no comparisons exist. I drop test scores in FIMS when developing Quality Index 2 as Hanushek and Kimko (2000) did, which does not affect the number of observations in Quality Index 2 because countries that took FIMS participated subsequent studies at least once.

[15] Standard errors are transferred in the same way as average test scores. Information on standard errors is not available in the Barro and Lee's quality dataset but can be obtained from various sources listed in Appendix A. For some countries where several education systems independently participated in an international comparative study, following Hanushek and Kim (1995), I simply take the arithmetic average for the standard error to get a single nominal standard error for that county,

measure of schooling quality for each country. Summary statistics for QL1 and QL2 are shown in Table 2. For comparison, summary statistics for schooling quantity is reported in Table 2 as well. To make sure that these two quality indexes pick up most of the information from each separate performance series, pairwise correlations have been estimated. The results show a high positive correlation between quality indexes and each performance series that they are based on.

Table 2 Summary statistics of quality indexes and schooling quantity

Quality and Quantity	Number of countries	Mean	Standard Deviation	Minimum	Maximum
QL1	52	46.56	9.46	23.44	60.74
QL2	52	48.62	10.22	21.65	63.69
S	48[a]	6.63	2.38	0.56	10.69

Notes: QL1 is Quality Index 1. QL2 is Quality Index 2. S is average years of adult schooling. [a] while 113 countries have schooling quantity data available, only 48 of them have participated in international tests at least once.

2.4 Other Data Sources

Real GDP per capita, population, the investment share of GDP and the ratio of total trade to GDP are obtained from Penn World Table 6.1 compiled by Summers, Heston and Aten (2002). The ratio of government consumption to GDP is from World Development Indicators 2001. School quantity measures are from Barro and Lee (2000) updated data set where data for educational attainment is presented at five-year intervals for the years 1960-2000. School inputs are acquired from the Barro and Lee (1996) data set[16] where school inputs data are reported at five-year intervals for the years 1960-1990.All data used in this paper are an updated version of the data in Hanushek and Kimko (2000).

[16] School input data can be obtained from the same website where data of test scores are provided.

3. Basic Results for Growth Effect of Quantity and Quality

Results for the baseline cross-country regressions are reported in Table 3, which is a replication of Table 2 in Hanushek and Kimko (2000) with updated data. The dependant variable is the annualized average growth rate of real GDP per capita for the years 1960-1995[17]. Whereas constructed quality indexes are available in 52 countries, only 39 of these countries have direct observations of real GDP growth rate and schooling quantity[18]. For comparability, only those 39 countries are considered in regressions even when observations for more than 39 countries are available in the case that quality indexes are not included. Column (1) of Table 3, the simplest model, includes initial real GDP per capita (Y60) and schooling quantity (S), where S is measured by the arithmetic average years of adult schooling for 1960, 1965, 1970, 1975, 1980, 1985, 1990, and 1995. The initial level of GDP is significantly and negatively related to growth rates, providing evidence of conditional convergence. Schooling quantity shows a positive and significant relationship with GDP growth rates. Annual population growth rate (GPOP) is added into the simplest model and results are shown in column (4). Nothing changes significantly with the inclusion of population growth rates, which is negatively and significantly associated with GDP growth.

[17] Hanushek and Kimko (2000) consider the time range from 1960 to 1990.

[18] For those 39 countries, QL1 has a mean 46.04 and a standard deviation of 10.02; QL2 has a mean 47.98 and a standard error of 11.02.

In columns (2) and (3) of Table 2 where quality indexes are added to the simplest regression, evidence of conditional convergence gets stronger and the coefficients of schooling quantity drop by about 50 percent in magnitude and lose significance. Schooling quality measures taking the place of quantity show a positive and significant impact on growth rates of GDP per capita. Based on the result in column (2), a one-standard-deviation increase in schooling quality is on average associated with an annual increase in per capita GDP of 1.12 (0.112*10.02) percentage points, holding schooling quantity and initial GDP constant. Likewise, a one-standard-deviation increase in schooling quantity is associated with an annual increase in GDP per capita of 0.52 (0.213*2.43[19]) percentage points, holding school quality and initial GDP constant. Columns (5) and (6) indicate population growth rates become insignificant when quality measures are included, whereas the inclusion of population growth does not make any qualitative changes in quality measures.

[19] For 39 countries, S has a mean 6.39 and a standard deviation of 2.43.

Table 3 Growth models with schooling quality

Dependent Variable: Annualized average growth rate in GDP per capita, 1960- 1995 (%)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial GDP per capita (Y60) ($1000)	-0.480 (0.109)	-0.400 (0.076)	-0.471 (0.087)	-0.509 (0.113)	-0.417 (0.077)	-0.488 (0.090)	-0.330 (0.062)	-0.390 (0.076)
Schooling quantity (S)	0.620 (0.198)	0.213 (0.177)	0.328 (0.195)	0.556 (0.209)	0.209 (0.178)	0.315 (0.199)	1.045 (0.470)	0.972 (0.572)
Annual population growth (%)(GPOP)				-0.450 (0.191)	-0.195 (0.201)	-0.260 (0.199)		
Schooling Quality (QL1)		0.112 (0.027)			0.107 (0.029)		0.214 (0.06)	
QL1*S							-0.021 (0.011)	
Schooling Quality (QL2)			0.090 (0.027)			0.084 (0.028)		0.164 (0.068)
QL2*S								-0.016 (0.012)
Constant	1.716 (0.918)	-1.287 (1.078)	-0.785 (1.097)	2.867 (1.077)	-0.639 (1.400)	0.043 (1.369)	-5.275 (2.559)	-3.876 (3.033)
R-square	0.43	0.66	0.61	0.47	0.67	0.62	0.71	0.64

Notes: Huber-White standard errors are in parentheses. Sample size is 39 countries. GPOP is average of annual population growth rate for the period 1960-1995.

R-squares for regressions (1) and (4) are boosted from around 40% to around 60% when QL1 or QL2 is controlled for, which indicates that schooling quality has a great contribution in explaining variations in GDP growth rates.

As the effects of schooling quality and quantity might depend on the level of the other, columns (7) and (8) of Table 3 add an interaction term between schooling quantity and quality to regressions (2) and (3), respectively. As showed in the table, the interaction term has a negative sign that is significant in the regression with QL1 but not in the regression with QL2. This implies that the marginal effects of schooling quality (quantity) are higher in countries with lower levels of schooling quantity (quality). However, implausible marginal effects of quality and quantity may occur for sample extremes based on the results in columns (7) and (8).

Regressions with logarithmic form of the right-hand side variables in the baseline models are also estimated, which produces identical qualitative results as Table 3.

4. Robustness Tests

This section examines if the estimated relationship between quality measures and economic growth is robust to the adding of some other explanatory variables and if it is sensitive to some East Asian countries.

4.1 Robustness to Other Explanatory Variables

Levine and Renelt (1992, p. 943) argue that "the cross-country statistical relationships between long-run average growth rates and almost every particular policy indicator considered by the profession are fragile: small alterations in the 'other' explanatory variables overturn past results". Thus, the robustness of the estimated relationship between quality measures and growth rates is tested by the inclusion of some other independent variables. Similar to Levine and Renelt (1992) and Hanushek and Kim (1995), I consider investment share of real GDP per capita (IVN), the ratio of total trade to real GDP per capita (TRD), and government consumption share of GDP (GOV).

Column (1) of Table 4 reproduces the result in column (2) of Table 3. Columns (2), (3) and (4) of Table 4 add each of the controlling variables into the first regression, respectively. Columns (5) and (6) include some pairs of those three variables. Quality measures remain significant with alterations in the conditioning information sets, although the magnitude of coefficients decreases in some degrees especially with the inclusion of the investment variable. The investment share of GDP also shows a robustly positive relationship with growth rates in GDP per capita, while the ratio of total trade to GDP is sensitive to the inclusion of the investment variable, which is consistent to the findings in Levine and Renelt (1992). Government consumption is negatively related with growth, but its significance depends on the model specifications. The substitution of QL1 with QL2 produces identical qualitative results which are not presented here to save space. The coefficients for schooling quantity are positive but insignificant through all regressions. Conditional convergence is evident in every regression.

Table 4 Robustness of quality effects on growth

Dependent Variable: Annualized average growth rate in GDP per capita, 1960-1995 (%)

	(1)	(2)	(3)	(4)	(5)	(6)
Initial GDP per capita Y60 ($1000)	-0.400	-0.368	-0.390	-0.347	-0.368	-0.302
	(0.076)	(0.073)	(0.078)	(0.074)	(0.074)	(0.064)
Schooling quantity (S)	0.213	0.162	0.236	0.209	0.170	0.149
	(0.177)	(0.143)	(0.178)	0.179	(0.144)	(0.141)
Schooling Quality (QL1)	0.112	0.073	0.103	0.106	0.074	0.053
	(0.027)	(0.026)	(0.024)	(0.028)	(0.027)	(0.018)
Investment/GDP		9.380			8.829	11.254
(INV)		(3.11)			(3.722)	(2.504)
Total trade/GDP			0.777		0.184	
(TRD)			(0.392)		(0.374)	
Government consumption/GDP				-5.794		-4.291
(GOV)				(2.343)		(2.313)
Constant	-1.287	-1.430	-1.480	-0.438	-1.471	-0.639
	(1.08)	(0.872)	(1.104)	(1.157)	(0.887)	(0.901)
R-square	0.66	0.76	0.69	0.65	0.76	0.79
No. of observations	39	39	39	38	39	38

Notes: Huber-White standard errors are in parentheses. INV is average ratio of investment to real GDP per capita for the period 1960-1995. TRD is average ratio of total trade to real GDP per capita for the period 1960-1995. GOV is average ratio of government consumption to GDP for the period 1960-1995.

4.2 Sensitivity to East Asian Countries

Many East Asian counties obtain higher scores in international tests than countries in other regions (see Table 5) and they also present high growth rates in real

GDP per capita. As a result, it is possible that the positive and significant relationship between quality measures based on test scores and growth is simply driven by some East Asian countries. Following Hanushek and Kimko (2000), I test the sensitivity of the results to East Asian countries by excluding different subsets of them. Results are reported in Table 6. Table 5 reports the mean of quality measures for the different subsets of East Asian countries considered in Table 6.

Table 5 Comparison of average quality indexes

	Full sample (39 countries)	Four Tigers[a]	High Performing[b]	Newly Industrialized[c]
Mean QL1	46.04	54.19	55.97	54.16
Mean QL2	47.98	54.10	56.31	54.54

Notes: [a]The "Four Tigers" are Hong Kong, Korea, Singapore and Taiwan.
High Performing are the Four Tigers plus Japan and China (Hanushek and Kimko,2000).
[c]Newly Industrialized are High Performing plus Indonesia, Malaysia and Thailand (Hanushek and Kimko, 2000) (Indonesia and Malaysia did not take any international tests.)

Column (1) of Table 6 presents the results from full sample with 39 countries which can be compared with results in column (2) where the Four Tigers are excluded and results in columns (3) and (4) where the High Performing and the Newly Industrialized countries are respectively excluded. Quality measures are always positively and significantly related to economic growth regardless of any subsets of East Asian countries being excluded, although the magnitude of coefficients for quality measures decreases compared to the full sample case. The substitution of QL1 with QL2 leaves nothing changed significantly.

Table 6 Sensitivity to East Asian countries

Dependent Variable: Annualized average growth rate in GDP per capita, 1960-1995 (%)

	(1) Full sample	(2) Excluding Four Tigers	(3) Excluding High Performing	(4) Excluding Newly Industrialized
Initial GDP per capita Y60 ($1000)	-0.400 (0.076)	-0.328 (0.080)	-0.358 (0.093)	-0.326 (0.086)
Schooling quantity (S)	0.213 (0.177)	0.234 (0.172)	0.205 (0.175)	0.216 (0.174)
Schooling Quality (QL1)	0.112 (0.027)	0.082 (0.021)	0.101 (0.027)	0.090 (0.025)
Constant	-1.287 (1.08)	-0.628 (1.065)	-1.063 (1.206)	-0.903 (1.173)
R-square	0.66	0.53	0.50	0.47
No. of observations	39	35	33	32

Note: Huber-White standard errors are in parentheses.

The robustness tests indicate that the estimated relationship between quality measures based on cognitive achievements and economic growth is robust to the alternations in the independent variables and is not driven by East Asian countries.

5. Comparison of Alternative Measures of School Quality

Previous studies of cross-country growth differences usually use school inputs as proxies for human-capital quality (e.g., Barro, 1991).The question is if quality indexes measured by cognitive skills are better measures for schooling quality than school inputs; or they simply capture different dimensions of human capital. The estimation results presented in Table 7 help us in answering this question.

Column (1) of Table 7 includes the pupil-teacher ratio in primary and in secondary schools (PT-pri and PT-sec) as proxies for school quality without the inclusion of quality indexes. Whereas the pupil-teacher ratio in secondary schools has an unexpected but insignificant association with economic growth, the coefficient for pupil-teacher ratio in primary schools has a negative sign, as expected, and is significant at 10% level. That means a country with a larger class size in primary schools has a lower growth rate in real GDP per capita on average. Education expenditure per pupil is considered independently as a measure of quality in column (2) of Table 7 without controlling for any quality indexes. The result shows a significant and positive association with growth. Columns (3) and (4) reproduce the results for the constructed quality indexes. More than 60% of variation in GDP growth rates can be explained by models with constructed quality indexes while R squares in models with schooling inputs are around 50% or less. When both school input measures and constructed quality measures are included (see column (5) and (6)), both pupil-teacher ratio in primary schools and per-pupil education expenditure lose significance. However, quality indexes are still significant as usual. These results imply that quality indexes based on cognitive skills are better measures for schooling quality than schooling inputs as they may capture more information on human capital than school inputs do.

Table 7 Alternative measures of school quality and growth

Dependent Variable: Annualized average growth rate in GDP per capita, 1960-1995 (%)

	School inputs		Quality Indexes		Combined input and index	
	(1)	(2)	(3)	(4)	(5)	(6)
Initial GDP per capita (Y60) ($1000)	-0.468 (0.102)	-0.594 (0.111)	-0.400 (0.076)	-0.471 (0.087)	-0.429 (0.084)	-0.521 (0.084)
Schooling quantity (S)	0.458 (0.154)	0.587 (0.201)	0.213 (0.177)	0.328 (0.195)	0.169 (0.149)	0.258 (0.15)
Pupil/teacher ratio in primary school (PT-pri)	-0.062 (0.034)				-0.031 (0.036)	-0.036 (0.039)
Pupil/teacher ratio in secondary school (PT-sec)	0.046 (0.073)				0.022 (0.054)	0.031 (0.056)
Per-pupil expenditure (PPE) ($1000)		0.333 (0.158)			0.073 (0.125)	0.173 (0.130)
Schooling Quality (QL1)			0.112 (0.027)		0.100 (0.034)	
Schooling Quality (QL2)				0.090 (0.027)		0.076 (0.033)
Constant	3.569 (1.32)	1.911 (0.93)	-1.287 (1.08)	-0.785 (1.097)	0.055 (1.288)	0.707 (1.431)
R-square	0.50	0.46	0.66	0.61	0.68	0.64

Notes: Huber-White standard errors are in parentheses. Sample size is 39 countries. PT-pri is average pupil-teacher ratio in primary schools for 1960, 1965, 1970, 1975, 1980, 1985, 1990. PT-sec is average pupil-teacher ratio in secondary schools for 1960, 1965, 1970, 1975, 1980, 1985, 1990. PPE is average education expenditure per pupil for 1960, 1965, 1970, 1975, 1980, 1985, 1990.

6. Causality – the Production of Schooling Quality

The results indicate that the constructed quality indexes are good measures of schooling quality which has a significant, positive and robust association with economic growth. However, there is nothing we can say so far about the causality of the estimated relationship between quality measures and growth. This issue is explored in this section.

Countries with higher rates of growth have the potential to invest more in school resources which in turn would improve quality of human capital if more school resources relate to better educational quality (Hanushek and Kimko, 2000). Hanushek and Kimko (2000) describe such relationships in the following functions:

(1) $g_i = X_i \beta + \gamma QL_i + \varepsilon_i$

(2) $R_i = W_i \delta + \eta g_i + v_i$

(3) $QL_i = Z_i \alpha + \pi R_i + \upsilon_i$

Equation (1) is the equation estimated in the previous sections, which relates country *i* 's growth rate (g_i) to some factors X_i and schooling quality (QL_i) in that country. Equation (2) shows that schooling resources (R_i) in country i are affected by growth rates (g_i) plus some other factors W_i. Equation (3) assumes schooling quality (QL_i) is a function of school resources along with some socioeconomic factors (Z_i) (e.g., family factors). Hanushek and Kimko (2000) argue that equations (2) and (3) reveal the fact that government can indirectly affect school outcomes by pursuing policies that influence school resources. If this is the case, the direct estimation of equation (1) would yield a biased coefficient estimate for schooling quality which partly captures the effect from growth to quality through the channel described in equations (2) and (3). As data limitations make it infeasible to estimate the whole system of equations, Hanushek and Kimko (2000) choose to estimate equation (3) directly by the approach of relating normalized individual country test scores on the test conducted at year *t* to the relevant family characteristics and school resources in year *t-1*. They find that school resources enter either with unexpected signs or with insignificantly expected signs, which leads to their conclusion that the estimated coefficients for quality measures reflect causal links from schooling quality to growth. Interestingly, Barro and Lee (2000) estimate the same equation as equation (3) by a different approach (SUR technique)[20] with the newly-compiled data for school resources, finding that school resources have a significant effect on international test scores. In what follows, in order to examine the causal link in cross-country growth models, I estimate the quality production function using the same dataset as Barro and Lee (2001) and employing the estimation approach of Hanushek and Kimko (2000).

I use the weighted average method described in section 2 to combine the normalized[21] test performance in mathematics and science by testing year. In effect, each testing year corresponds to one specific international comparative study (see Table 1). Seven combined test series are developed and I call them test1965, test1970, test1980, test1985, test1988, test1991 and test1995. The series test1965, for example, is the weighted average of test performance for 13-year-old students and students at final year

[20] A system of 13 equations are estimated by the seemingly-unrelated-regression (SUR) technique with the dependent variables as the disaggregated scores on international tests in mathematics, science and reading in various years for 10 and 14-year-old students. Their basic regressions assume same slope coefficients for the independent variables but allow for different intercepts. Two variant regressions that allow for different slope coefficients in the equations for each age group and for each subject areas are also estimated by Barro and Lee (2001).

[21] Each test performance is normalized to a mean of 50.

of secondary education. Each combined test series has a considerably high correlation, usually over 0.8, with the underlying separated test performance of students. Then, I relate each test at year *t* to the relevant family characterizes (Z) and school resources (R) at year *t-1*. Thus, the quality production function estimated is:

$$QL_{it} = Z_{it-1}\,\alpha + \pi\, R_{it-1} + \upsilon_{it.}$$

As described in equation (2), school resource of country i at time *t-1* (R_{it-1}) is determined by growth of country i at year *t-1* (g_{it-1}), other factor W_{it-1}, and the disturbance term $\nu_{it-1.}$ Estimation of the quality production function would yield consistent coefficient estimates if ν_{it-1} is uncorrelated with υ_{it} and the mean of υ_{it} conditional on Z_{it-1} and R_{it-1} is zero.The dating of independent variables is presented in Appendix B.

Results from panel regressions are presented in Table 8. Although there are 125 direct observations of combined test scores, only up to 113 observations also have right-hand side data available. Regressions (1), (2) and (3) of Table 8 consider, respectively, education expenditure per pupil (PPE_{t-1}) and pupil-teacher ratio in primary schools (PT-pri $_{t-1}$) and average teacher salary in primary schools ($SALARP_{t-1}$) as characteristics of school system. Expenditure per pupil is significant but with an unexpected sign. Pupil-teacher ratio in primary schools also has an unexpected sign but it is insignificant. Although teacher salary presents a positive effect on test scores, it is insignificant. These results are consistent with those in Hanushek and Kimko (2000).

Table 8 The determinants of schooling quality

Dependent variable: Combined test performance at test year *t*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adult education (S_{t-1})	2.02	1.62	1.18	2.05	1.76	1.55	2.16	2.25
	(0.63)	(0.48)	(0.55)	(0.57)	(0.40)	(0.49)	(0.65)	(0.55)
PT-pri_{t-1} [a]		0.03			-0.11		-0.12	-0.25
		(0.15)			(0.09)		(0.16)	(0.13)
Per pupil expenditure (PPEt-1) ($1000)	-1.22			-0.60			-3.49	-3.03
	(0.54)			(0.50)			(0.96)	(0.83)
SALAR-pri_{t-1} [b] ($1000)			0.06			0.10	0.45	0.42
			(0.07)			(0.07)	(0.12)	(0.10)
Annual population growth (%) ($GPOP_{t-1}$)	-2.02	-1.54	-0.75	-3.06	-2.59	-2.28	-1.02	-1.83
	(1.18)	(1.41)	(1.25)	(1.48)	(1.65)	(1.68)	(1.40)	(1.67)
East Asian dummy				10.30	12.05	10.88		10.32
				(0.45)	(0.31)	(2.55)		(2.48)
Constant	41.44	39.41	41.75	38.59	40.61	37.80	40.57	41.20
	(3.91)	(5.25)	(4.12)	(3.63)	(4.27)	(3.64)	(4.71)	(3.76)
R-square	0.21	0.21	0.14	0.38	0.41	0.34	0.27	0.43
No. of observations	107	113	104	107	113	104	104	104

Notes: Huber-White standard errors are in parentheses.
[a]PT-pri represents Pupil teacher ratio in primary schools.
[b]SALAR-pri represents average teacher salary in primary schools.

Columns (4), (5) and (6) of Table 8 report the results for adding an East Asian country indicator to each of the first three regressions. The expenditure variable still has a negative effect on test scores but becomes insignificant and pupil-teacher ratio in primary schools changes to the expected sign although it is still insignificant. Contrary to the finding in Hanushek and Kimko (2000), my results indicate that the perverse impacts of pupil-teacher ratios might be an "artifact of the large class sizes" (p.1193) in most East Asian countries. Nothing significantly changes for teacher salary with the region indicator included.

Most interestingly, when all three school resource variables are included at the same time (see column (7)), average teacher salary has a significantly positive effect on test performance and pupil-teacher ratio presents an insignificantly negative sign. When an East Asian country indicator is included in column (8), pupil-teacher ratio in primary schools changes to be significant at 10% level with a negative sign. However, education expenditure per pupil shows a significantly negative effect on scores in the last two regressions. These findings indicate that holding everything else constant including education expenditure per pupil and class size, increases in teacher salary in primary schools (with less expenditure on other school inputs) can improve student performance in international tests. This result is consistent with the findings in Barro and Lee (2000).

Throughout the regressions in Table 8, the socioeconomic variables --- general education level of parents (S_{t-1}) and growth rate of population ($GPOP_{t-1}$) --- always have the expected signs. Among them, parental education presents a positive and significant effect on student test scores in all specifications. The effect of growth rate in population is more sensitive to the model specifications.

Replacing pupil-teacher ratio in primary schools with pupil-teacher ratio in secondary schools and per pupil expenditure with the ratio of per pupil expenditure to GDP does not change results significantly. Results also hold for the inclusion of other socioeconomic variables (e.g., average income of parents).

In short, my results based on updated data differ from those in Hanushek and Kimko (2000) in two ways. First, the perverse signs of pupil-teacher ratio could be driven by East Asian countries which usually have large class sizes and obtain high scores in international tests as well. Second, the effects of school resources on test scores are sensitive to model specifications. In some model specifications, school resources show a positive effect on test scores, which probably occurs through increasing teacher salary in primary schools while holding per-pupil education expenditure and pupil-teach ratio in primary schools constant. Thus, reverse causality might be a concern for the estimated growth effect of constructed quality measures in cross-section regressions.

7. Conclusion

The literature on growth effects of education tends to focus on quantity measures of schooling, leading to some findings that cast doubt on the role of human capital in long run economic growth. Hanushek and Kimko (2000) emphasize the importance of labor-force quality in determining growth. This paper follows Hanushek and Kimko (2000) by investigating the role of schooling quality in economic growth. Consistent with Hanushek and Kimko (2000), I find out that quality measures developed from expanded international comparative studies in mathematics and science have a strong positive relationship with cross-country differences in real GDP growth rates. This relationship seems to be robust to alternations in the conditioning variable sets and is not attributed to some East Asian countries that have higher performance in both economic growth and international tests. Comparison of alternative measures of schooling quality provides further support that constructed quality measures are better measures of quality than school inputs. However, education production regressions provide some evidence against findings in Hanushek and Kimko (2000) and supporting findings in Barro and Lee (2000) that reverse causality can not be ruled out. That is to say, the estimated coefficients for quality measures in cross-sectional regressions might partly reflect the effect of growth on schooling quality through governmental investment in school resources.

Another finding that is worth mention is that schooling quantity loses its significance almost whenever quality indexes are considered. Thus, one conclusion that can be made with more confidence is schooling quality measured by cognitive skills is a more important factor, at least in the statistical sense, than schooling quantity in determining a country's economic growth rate. This suggests that policy makers should not only pay their attention to the increase of educational quantity but attach more importance to the improvement of educational quality.

Appendices

Appendix A.
Data Sources for Standard Errors

Test	Subjects	Age Groups	Sources of Standard Errors
FIMS	Math	13, Final sec[a].	Medrich and Griffith (1992)
FISS	Science	10, 14, Final sec.	Medrich and Griffith (1992)
SIMS	Math	13, Final sec.	Medrich and Griffith (1992)
SISS	Science	10, 14, Final sec.	Medrich and Griffith (1992)
First IAEP Study	Math	13	Medrich and Griffith (1992)
	Science	13	Medrich and Griffith (1992)
Second IAEP Study	Math	9, 13	Digest of Educational Statistics (1992)
	Science	9, 13	Digest of Educational Statistics (1992)
TIMSS	Math	9	Mullis et al. (1997)
	Math	13	Beaton et al. (1997)
	Math	Final sec.	Mullis et al. (1998)
	Science	9	Martin et al. (1997)
	Science	13	Beaton et al. (1997)
	Science	Final sec.	Mullis et al. (1998)

Note: [a]Final sec. denotes the final year of secondary education.

Appendix B.
Dating of Independent Variables for the Education Production Models

	IEA math1 Test1965	IEA science1 Test1970	IEA math2 Test1980	IEA science2 Test1985	IAEP study1 Test1988	IAEP study2 Test1990	IEA TIMSS Test1995
S_{t-1}	1960	1960-1965	1965-1975	1970-1980	1975-1985	1975-1985	1980-1990
PT-pri_{t-1}	1960	1960-1965	1965-1975	1970-1980	1975-1985	1975-1985	1980-1990
PPE_{t-1}	1960	1960-1965	1965-1975	1970-1980	1975-1985	1975-1985	1980-1990
SALAR-pri_{t-1}	1960	1960-1965	1965-1975	1970-1980	1975-1985	1975-1985	1980-1990
$GPOP_{t-1}$	1960-1964	1960-1969	1965-1974	1970-1979	1975-1984	1975-1984	1980-1989

Notes: S represents average years of adult schooling. PT-pri represents pupil teacher ratio in primary schools. PPE is per-pupil expenditure. SALAR-pri represents average teacher salary in primary schools. GPOP is population growth rates.
Source: adapted from Hanushek and Kimko (2000)

Appendix C. Data for Quality Indexes

	Country	QL1	QL2		Country	QL1	QL2
1	Australia	50.796	55.066	27	Korea	57.876	57.417
2	Austria	53.832	56.884	28	Latvia	45.13	42.463
3	Belgium	54.569	53.504	29	Lithuania	42.951	49.392
4	Brazil	33.397	36.205	30	Luxembourg	39.786	44.327
5	Bulgaria	56.3	56.278	31	Mozambique	23.8	27.635
6	Canada	50.713	53.405	32	Netherlands	55.246	57.291
7	Chile	27.655	25.537	33	New Zealand	52.998	59.207
8	China	58.33	63.691	34	Nigeria	35.062	39.063
9	Colombia	30.575	31.462	35	Norway	49.781	53.82
10	Cyprus	42.463	45.719	36	Philippines	34.542	33.064
11	Czech Republic	56.174	53.433	37	Poland	48.58	49.13
12	Denmark	49.231	56.656	38	Portugal	41.83	40.708
13	Finland	52.647	55.373	39	Romania	44.646	45.371
14	France	50.946	59.584	40	Russian Federation	51.179	55.772
15	Germany	53.902	58.391	41	Singapore	51.358	50.763
16	Greece	42.618	40.062	42	Slovak Republic	55.041	61.65
17	Hong Kong	51.731	52.405	43	South Africa	30.022	34.388
18	Hungary	54.785	56.35	44	Spain	48.581	50.227
19	Iceland	50.005	56.337	45	Swaziland	35.587	39.648
20	India	23.441	21.645	46	Sweden	51.885	56.813
21	Iran, I.R. of	28.644	26.859	47	Switzerland	53.346	59.103
22	Ireland	48.714	49.109	48	Taiwan	55.805	55.83
23	Israel	53.574	52.77	49	Thailand	43.261	43.909
24	Italy	49.12	51.621	50	United Kingdom	50.876	52.094
25	Japan	60.739	57.774	51	United States	47.357	50.206
26	Jordan	38.778	41.801	52	Yugoslavia	51.1	50.889

Reference List

Barro, R. J. (1991). Economic growth in a gross section of countries. *Quarterly Journal of Economics, 106*(2), 407-443.

Barro, R. J., & Lee, J.-W. (1996). International measures of schooling years and schooling quality. *American Economic Review, 73*, 928-946

Barro, R. J., & Lee, J.-W. (2000). International data on educational attainment: updates and implications. Working paper, Center for International Development, Harvard University.

Barro, R. J., & Lee, J.-W. (2001). Schooling quality in a cross-section of countries. *Economica, 68*, 465-488

Beaton, A. E., Martin, M. O., Mullis, I., Gonzalez, E. J., Kelly, D. L., & Smith T. A. (1997). *Mathematics Achievement in the Middles School Years: IEA's Third International Mathematics and Science.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Beaton, A. E., Mullis, I., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith T. A. (1997). *Science Achievement in the Middles School Years: IEA's Third International Mathematics and Science.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development: evidence from aggregate cross country data. *Journal of Monetary Economics, 34*(2), 143-174.

Hanushek, E. A., & Kim, D. (1995). Schooling, labor force quality, and economic growth. National Bureau of Economic Research Working Paper 5399

Hanushek, E. A., & Kimko, D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review, 90*(5), 1184-1208.

Heston, A., Summers R., & Aten, B. (2002). Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP).

Islam N. (1995). Growth empirics: a panel data approach. *Quarterly Journal of Economics, 110*(4), 1127-1170

Kruger, A. B., & Lindahl, M. (2001). Education for growth: why and for whom. *Journal of Economic Literature. 39*(4), 1101-1136.

Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review 82*(4), 942-963.

Mankiw, N. G., Romer, D., & Weil, D. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics, 107*(2), 402-437.

Medrich, E. A., & Griffith, J. E. (1992). *International Mathematics and Science Assessment: What Have we Learned?* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Martin, M. O., Mullis, I., Beaton, A. E., Gonzalez, E. J., Smith, T. A, & .Kelly, D. L. (1998). *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Mullis, I., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Mullis, I., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics and Science Achievement in the Final Years of Secondary School: IEA's Third International Mathematics and Science.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Nelson, R. R., & Phelps, E. (1966). Investment in humans, technology diffusion, and economic growth. *American Economic Review, 56*(2), 69-75.

Pritchett, L. (1996). Where has all the education gone? Mimeo. World Bank Policy Research Department Working Paper 1581.

Rebelo, S. (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy, 99*(3), 500-521.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy, 94*(5), 1002–1037

Romer, P. M. (1990a). Endogenous technological change. *Journal of Political Economy 98*(5), S71-102

U.S. Department of Education (1992). *Digest of Educational Statistics*. Washington DC: National Center for Educational Statistics.

U.S. Department of Education (1997). *Digest of Educational Statistics*. Washington DC: National Center for Educational Statistics.

U.S. Department of Education (1994). *The Condition of Education 1994.* Washington, DC: National Center for Education Statistics.