Quality-Consistent Estimates of International Schooling and Skill Gradients

Eric A. Hanushek
Stanford University

Lei Zhang
Clemson University

Mincer wage equations focus on the earnings premium associated with additional schooling for a cross section of individuals of different ages but generally fail to account for changes in education quality over time. More fundamentally, school attainment is an inadequate proxy of individual skills, when both family inputs and ability affect cognitive skills. We combine quality-adjusted measures of schooling and international literacy test information to estimate skill gradients for 13 countries. The premiums to quality-adjusted education are considerably higher than the traditional Mincer estimate for most countries, but this bias is more than offset by consideration of other factors affecting skills and earnings.

I. Introduction

School attainment has been the central focus of many policy discussions around the world since the influential work of Mincer (1970, 1974) identified schooling as the prime proxy for human capital and individual labor market skills. His innovative empirical approach to investigating earnings determination across individuals has fueled continuing investigations of national and international differences in the returns to schooling (e.g., Psacharopoulos 1994; Harmon, Oosterbeek, and Walker 2003; Psacharopoulos and Patrinos 2004). But a parallel literature has highlighted potential problems with estimation and interpretation of the standard Mincerian model, centering largely on whether the schooling coefficient in a cross-sectional earnings analysis should really be viewed as the rate of return to a year of school. We find, however, that even dealing with the issues in these critiques does not resolve some more fundamental problems. Neglect of school quality and other determinants of skills significantly confounds this estimation and affects interpretations of standard labor market analyses both within and across

This research has been supported by the Packard Humanities Institute. We thank Statistics Canada and Yvan Clermont, Richard Desjardins, Scott Murray, and Albert Tuijnman for helping to provide us with the data.

[*Journal of Human Capital,* 2009, vol. 3, no. 2]

countries. Our analysis indicates that fundamental issues of the measurement of individual skills are centrally important but have been largely ignored.

We focus on two aspects of the measurement of skills that are extremely important—particularly in an international context.[1] First, if the quality of schooling obtained differs across time, the estimated average wage return to schooling of different qualities may over- or underestimate the earnings differential from varying school attainment depending on how schooling quality has changed over time. Consider, for example, the standard cross-sectional estimation of Mincer earnings models that relate an individual's school attainment to earnings. A fundamental presumption is that, say, the cross-sectional earnings of a 45-year-old high school graduate are a good indication of what a 25-year-old graduate can expect in 20 years.[2] This would not be the case if high school graduation implied different skills for individuals of different ages. Second, because school attainment is an imperfect proxy for individual skills, the estimated schooling differential does not necessarily reflect the causal effect of changing school attainment. The importance of these two issues depends not only on variations in the quality of schooling but also on how nonschool factors affect skills. Ignoring these issues turns out to be especially important for international comparison of the returns to skills.

We first construct a quality-adjusted years of schooling measure for individuals of 13 countries separately. This quality adjustment standardizes schooling obtained at different points in time on the basis of the relative contributions of schooling to cognitive skills. For almost all countries in our sample, our analysis suggests that the contributions of additional schooling to literacy skills are higher for more recent cohorts. This is consistent with the average quality of schooling improving over time and in general implies that the simple Mincer schooling parameter underestimates the economic value of an additional year of schooling today. Compared to other countries, however, the adjustment of quality of education in the United States is less important, reinforcing the general finding that school quality in the United States has been relatively stable for several decades (see Hanushek 2003; National Center for Education Statistics 2005).

[1] Past attention to how good schooling is as an indicator of skills has focused on measurement errors. First, misreporting of school attainment on surveys threatens earnings estimation. From their sample of twins, Ashenfelter and Krueger (1994), e.g., conclude that measurement error is more important than selection problems in the estimation of the returns to schooling. Second, there is the strong suggestion that a general equivalency diploma is not the same as a regular high school diploma even though the two are frequently indistinguishable in available survey data (see Cameron and Heckman 1993; Tyler, Murnane, and Willett 2000).

[2] Prior attention to the age comparisons has focused on the possibility that patterns of technological change and productivity growth could systematically alter the future labor market returns to schooling (Katz and Murphy 1992; Murphy and Welch 1992). This also enters into the analysis of Heckman, Lochner, and Todd (2006).

Our estimates of the returns to skill show vividly how the standard estimation approaches to models of wage determination misstate the skill gradients within individual countries. They also provide a consistent way to compare the outcomes of labor markets in a range of countries as a first step in understanding how attributes of different economies affect the valuation of individual skills.[3]

The paper is organized as follows. Section II describes the relationship between an underlying skills model and the frequently estimated Mincer schooling model. Section III describes the underlying international data. Section IV sets up and estimates the empirical model of the quality of schooling at different time periods and develops a quality-adjusted years of schooling measure. Section V presents the estimation of returns to quality-adjusted years of schooling and to cognitive skills. In both empirical sections, we discuss the similarities and differences across countries.

II. The Basic Skills Model

Perhaps the simplest human capital model of wage determination is that wages are proportional to skills as in

$$w_i = \lambda H_i + \eta_i, \tag{1}$$

where w_i is the wage of individual i, H_i is skill or human capital, η_i is a random error, and λ is the wage return to skill. This simple formulation, however, belies two complexities: How are skills measured? And where do skills come from?

The innovation of Jacob Mincer was putting structure on equation (1) that provided both a simple, easily estimated empirical specification and an interpretation linking the estimates to meaningful investment parameters. In the most frequently employed variant, log earnings are associated with years of schooling for a cross section of individuals of varying ages. Potential experience (i.e., time out of school) and potential experience squared are typically added to reflect on-the-job training. Under a series of assumptions, the schooling coefficient in a log earnings model can be interpreted as the internal rate of return to schooling.

One central concern running through a large literature has been that higher-ability individuals may systematically choose more schooling, leading to an upward bias in the estimated return to schooling (see, e.g., the review and critique in Card [1999]). This line of analysis takes

[3] As Heckman, Lochner, and Todd (2006) demonstrate, going from estimates of skill gradients to the calculation of internal rates of return is not entirely direct. In an international context, this translation would be especially difficult. Moreover, this analysis introduces an additional complication because a key assumption in the Mincer interpretation is that the full cost of a year of schooling is the opportunity cost of not being in the labor market. However, our quality-adjusted schooling years may involve more or less than a year out of the labor market.

the perspective, generally without mention, that the relevant human capital measure is school attainment (i.e., $H_i \equiv S_i$). It then considers how mismeasurement of S or other factors that might affect costs of schooling or ability can influence estimates of λ.[4] Interestingly, most of the alternative estimates of λ yield results that are close to the ordinary least squares estimates, suggesting that these concerns are not such a great empirical problem.

A second class of concerns focuses specifically on the interpretation of the estimated schooling parameter from the basic Mincer specification and is more telling. The Mincer earnings function is frequently interpreted as providing direct estimates of internal rates of return to school attainment, and these can then be used in policy analyses about school investments. This interpretation, for example, underlies the extensive international comparisons (Psacharopoulos 1994; Harmon et al. 2003; Psacharopoulos and Patrinos 2004). The empirical analyses of Heckman, Lochner, and Todd (2006, 2008), however, demonstrate quite conclusively that the necessary assumptions about tuition costs, taxes, and functional form underlying this interpretation are not valid within U.S. data and that their failure leads to substantial bias compared to a properly calculated internal rate of return. This leads us to concentrate on the wage gradient associated with schooling and skills, which is a building block for calculating internal rates of return but is not equivalent.

Our analysis, however, steps back and takes a perspective different from that of most previous analyses. Our starting point is simply that human capital is about skills of workers that are rewarded in the labor market. School is one place where these skills are developed, but it is not the only place. Family, friends, neighbors, individual ability, and more contribute to an individual's skills. Because these other factors will in general be correlated with the school attainment of an individual, estimates of the labor market payoff to schooling will be biased by not considering them.

The concentration on school attainment in earnings functions has been an expedient because census information and labor market surveys typically have data on the three key elements: earnings, school attainment, and age. These data sources do not, however, tend to have information on variations in skills within schooling categories or on other sources of skill development.

Following Hanushek and Woessmann (2008), we presume as a first approximation that the relevant skills are cognitive skills that are measured, albeit imperfectly, by standardized achievement tests. While some

[4] A subset of the approaches bears a similarity to the analysis here by including measures of parental education as controls in the earnings models, although the choice of what controls to use or whether to use them as controls or instruments is generally unclear from the comparisons (see Card 1999).

past analyses have incorporated test scores into earnings analyses, they have generally taken them to be measures of "ability," suggesting that they are fixed early in life. Thus, ability affects school attainment (and labor market outcomes) but is itself unaffected by schooling. On the other hand, our perspective—matching much of the educational policy discussion—is that schools directly affect cognitive skills. This perspective then suggests that the broad literature on educational production functions is relevant for considering the formation of individual skills and thus is also important for consideration of earnings models (e.g., Hanushek 2002).

The linkage of skills to educational policy suggests that school quality may be an important element in earnings determination. Specifically, a year of high-quality schooling would be expected to impart more skills than a year of low-quality schooling. Such a possibility is particularly relevant for earnings analyses that rely on the cross-sectional variation in earnings across different cohorts. If there have been secular changes in the quality of schooling, measured school attainment for people educated in different eras with different quality of schools will not be equivalent in terms of skills.

Accounting for secular changes in school quality has been difficult within most available cross-sectional or panel data sets because there are no data that track quality with any precision. While some attempts rely on changes in measurable inputs—such as spending or pupil-teacher ratios—the uncertain verification of these measures of quality has led to limited acceptance.[5] Here we start with cognitive skills as a measure of the outcome of schooling. If providing cognitive skills is the primary objective of schools, we can use external information about student cognitive skills for individuals educated during different periods to provide information on changing school quality.[6]

By employing a unique data set that provides information on cognitive skills, schooling, and earnings across people of different ages, we are able to estimate quality-adjusted schooling for individuals in different cohorts and different countries. We can then estimate wage gradients with respect to quality-consistent schooling and skills and compare these to commonly available estimates.

III. The IALS Data

The primary data source is the International Adult Literacy Survey (IALS), conducted by the Organization for Economic Cooperation and Development (OECD). Twenty-three countries and regions participated

[5] This debate can be traced through Card and Krueger (1992), Hanushek, Rivkin, and Taylor (1996), and Heckman, Layne-Farrar, and Todd (1996); see also Hanushek (2003).

[6] The potential role of noncognitive skills, as discussed variously in Bowles, Gintis, and Osborne (2001) or Heckman, Stixrud, and Urzua (2006), is discussed below.

in one of three different waves of surveys conducted in 1994, 1996, and 1998.[7] The IALS is designed to compare individual literacy and numeracy skills within and across countries. (Throughout this paper we will follow the convention in the IALS survey of referring to the collection of separate tests of domains of basic skills related to the ability to use different kinds of information simply as "literacy.") A representative sample of adults between 16 and 65 years of age in each country were given a series of assessments of cognitive skills in the language of their country of residence. The literacy skill measures were supplemented by variables measuring other individual characteristics, such as country of origin, age, education, employment, earnings, and parents' education attainment.

Note that the oldest sampled individuals were born around 1930, and the youngest in the sample (16-year-olds) were born around 1980. The sample, heavily weighted toward European countries, thus has significant numbers attending school around World War II and during postwar reconstruction—suggesting that school quality may differ significantly for individuals with the same attainment but educated at different times within each country.

The IALS provides measurement of cognitive skills in three different areas. *Prose literacy* measures the knowledge and skills needed to understand and use information from texts including editorials, news stories, poems, and fiction. *Document literacy* measures the knowledge and skills required to locate and use information contained in various formats, including job applications, payroll forms, transportation schedules, maps, tables, and graphics. *Quantitative literacy* measures the knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials, such as balancing a checkbook, calculating a tip, completing an order form, or determining the amount of interest on a loan from an advertisement. The literacy scores range on a scale from 0 to 500 points for each area. Since the literacy scores are highly correlated with each other, we use the average of the scores in the analysis.

Table 1 provides summary statistics for the participating countries; the 13 countries in bold included continuous earnings measures and are included in the subsequent labor market analysis.[8] Sample sizes range from 2,062 in Germany to 5,660 in Canada. On the literacy tests, individuals score an average of 267 points with a standard deviation of 62 points. Sweden and Norway have the highest average, and Chile is at the bottom. Columns 3–6 show the considerable variation in not only average school attainment but also the distribution. For example, Chile

[7] A technical description of the survey and data can be found in Murray, Kirsch, and Jenkins (1997). The data are available from Statistics Canada (http://www.statcan.ca/english/freepub/89-588-XIE/about.htm).

[8] As discussed below, Canada, Slovenia, and the Italian region of Switzerland have continuous wage measures but are missing other crucial data needed for the full estimation.

TABLE 1
Descriptive Statistics for IALS Sample Countries

Country (Survey Date)	Observations (1)	Average Literacy Score (2)	Standard Deviation (3)	Years of Schooling (4)	Less than Upper Secondary (%) (5)	Tertiary and Above (%) (6)
All	64,196	267.4	62.2	12.1	38.8	25.2
Belgium (1996)	2,261	277.3	54.9	12.1	42.5	22.2
Canada (1994)	5,660	270.9	67.5	12.0	36.1	31.5
Chile (1998)	3,583	216.5	54.9	9.7	59.3	17.8
Czech Republic (1998)	3,132	283.5	46.0	12.4	58.9	11.1
Denmark (1998)	3,026	289.1	40.3	12.5	29.3	24.2
Finland (1998)	2,928	288.0	47.7	12.2	33.4	18.7
Germany (1994)	2,062	284.7	42.3	11.4	59.6	15.2
Great Britain (1996)	3,811	267.2	62.3	12.1	60.8	19.2
Hungary (1998)	2,593	253.8	48.0	11.7	29.4	16.8
Ireland (1994)	2,423	263.2	56.9	10.4	53.3	17.6
Italy (1998)	2,974	243.7	60.5	10.4	56.7	8.1
Netherlands (1994)	3,090	281.4	46.9	12.4	45.4	17.2
New Zealand (1996)	4,223	271.7	54.3	12.0	49.3	25.4
Northern Ireland (1996)	2,907	265.0	61.9	12.3	67.5	15.4
Norway (1998)	3,307	294.1	44.8	11.7	13.3	25.1
Poland (1994)	3,000	229.4	64.1	11.0	63.1	14.4
Slovenia (1998)	2,972	234.8	62.2	11.1	36.7	14.0
Sweden (1994)	3,038	293.4	55.1	11.0	36.1	22.5
Switzerland (French, German) (1994)	2,843	271.1	57.1	12.3	19.4	19.5
Switzerland (Italian) (1998)	1,302	269.9	51.1	11.7	37.8	11.7
United States (1994)	3,061	272.2	65.4	13.2	18.6	37.2

Note.—Bold indicates countries used in subsequent earnings analysis.

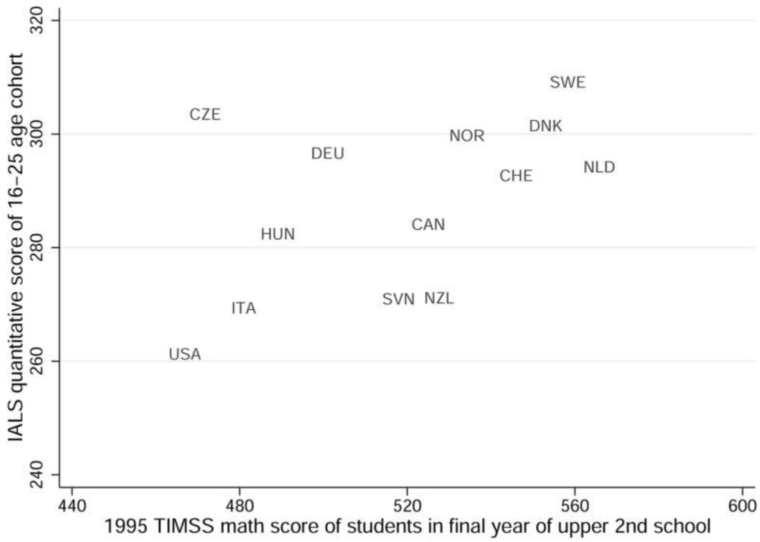

Figure 1.—IALS quantitative score and TIMSS math score. CAN stands for Canada, CHE for Switzerland, CZE for Czech Republic, DEU for Germany, DNK for Denmark, HUN for Hungary, ITA for Italy, NLD for the Netherlands, NOR for Norway, NZL for New Zealand, SVN for Slovenia, and SWE for Sweden.

and the Netherlands have similar completion of tertiary schooling (around 16 percent), but their average attainment differs by almost 3 years.

The literacy tests are designed to measure basic skills needed to participate fully in modern society, and it is useful to put these literacy test scores in the perspective of cognitive tests requiring deeper content knowledge and analytical skills. We compare the quantitative IALS score of individuals between 16 and 25 years of age to the 1995 TIMSS math score of students at the final year of upper secondary education, who are between 17 and 20 years of age.[9] Figure 1 illustrates the relationship between the TIMSS math score and young adults' quantitative literacy score. Thirteen countries are included in both the IALS and the TIMSS. The correlation between the average country scores is .73 and is significantly different from zero.[10] Thus, the literacy scores appear to be a reasonable index of general levels of skills.

[9] The Third International Mathematics and Science Study, or TIMSS, conducted in 1995 involved the participation of 40 countries and followed two prior test development cycles for math and for science. It is commonly accepted as a valid test for differences in math skills and includes a variety of high-level items covering calculus, probability and statistics, and geometry (see http://timss.bc.edu/timss1995.html). With testing in 1999 and after, TIMSS was renamed the Trends in International Mathematics and Science Study.

[10] The correlation is calculated for 12 countries: Czech Republic appears to be an outlier in the scatter plot. The same relationship holds when we restrict the IALS sample to the same age group as the TIMSS, but we lose Canada because it does not have an age measure. The same relationship holds for males and females separately.

IV. Calibrating School Quality

The goal of the empirical analysis is to provide estimates of how earnings vary both with schooling of a given quality and with cognitive skills. Conceptually, one would like to follow groups of individuals with differing investments in human capital over their entire careers and observe how earnings evolve and differ. This conceptual best may not, however, be ideal because one would not like to be restricted just to evaluating human capital investments made multiple decades earlier. An appealing analytical solution, laid out clearly in Mincer (1970, 1974), is to use data about otherwise similar individuals who provide investment-earnings observations at different points in the life cycle. The key question, one that has driven much of the subsequent research, is when are individuals "otherwise similar"?

Our focus is ensuring that individuals are comparable in terms of both school quality and cognitive skills. In this section, we estimate quality indices for schooling received at different time periods and adjust years of schooling with these quality indices relative to a base cohort. In the following section, we use the quality-adjusted years of schooling in a Mincer wage regression designed to estimate the impact on earnings of schooling and cognitive skills for this base cohort.

A. Identifying Intertemporal Changes in School Quality

Our modeling of school quality follows from the general analyses of educational production. The analyses of educational production typically relate measures of student achievement to underlying inputs of families, peers, and schools. Relying on cross-sectional or panel data on achievement, it focuses on various descriptions of how outcomes vary across students in a given grade of school at a specific time. Our focus here, however, slices the problem in a different way. We investigate how varying amounts of schooling affect cognitive skills at different points in time, holding constant the influence of families, ability, and other inputs to skills. This analysis is designed to indicate how overall school quality has changed within each country, so that it is possible to compare a year of schooling obtained at two different times.

We define a year of quality-equivalent schooling as one that produces the same increment in average cognitive skills, regardless of the time period in which it was received.[11] The objective is to identify the marginal impact of schooling on skills, holding constant other influences on skills. We operationalize this by relating our individual literacy scores to school attainment and other achievement inputs for individuals within each

[11] As will be apparent, our focus is the operation of schools and labor markets within individual countries, and we make no attempt to estimate quality-equivalent years of schooling across countries. For a related consideration of school attainment and cognitive skills in an international context, see Hanushek and Woessmann (2008).

country and then estimating the implicit quality adjustment for each time period and country from this.

Consider the determinants of literacy (L) for individual i in country k. We divide individuals into five cohorts (c) aged 16–25, 26–35, 36–45, 46–55, and 56–65 at the time of the IALS survey (which will correspond to distinct birth years and periods of schooling).[12] For each country, we then estimate an intertemporal production function such as

$$L_{ikc} = \beta_k(q_{kc}S_{ikc}) + X_{ik}\gamma_k + \varepsilon_{ikc}, \tag{2}$$

where L_{ikc} is the literacy score of individual i of country k and cohort c; S_{ikc} is the years of schooling of individual i of country k and cohort c; q_{kc} is the average quality of a year of schooling obtained by cohort c in country k; X_{ik} is a vector of control variables relevant to individual i with the country-specific impacts γ_k; and ε_{ikc} is a stochastic error term.[13] The term β_k measures the marginal contribution to the literacy score of the schooling of cohort c in country k, where each schooling year has quality q_{kc}.

Our objective is to estimate q_{kc} so that we can then develop a quality-equivalent schooling measure for each individual that can be applied to the analysis of earnings determination. Specifically, given estimated quality, we can then estimate quality-equivalent schooling for each individual as

$$\tilde{S}_{ikc} = \hat{q}_{kc}S_{ikc}. \tag{3}$$

The estimation and interpretation of q_{kc} is complicated by the fact that individual literacy reflects not only the education attainment of an individual but also other factors that vary across individuals and over time. We approach this by conditioning on other major influences on literacy over time (X).

First, average years of schooling have been constantly increasing for virtually every country over the past several decades.[14] Associated with this improvement is the concern that the school and college selectivity has gone down over time. In other words, if school continuation is related to ability, people with lower innate ability on average have been

[12] In a sensitivity analysis in which the sample is limited to individuals over 25 because of the concern that many of the 16–25 age group may still be in school, introducing sample selection biases in both the school quality analysis and the earnings analysis, we obtain almost identical results.

[13] One potential complication is that students may decide on staying in or dropping out of school in part on the basis of the quality of a school, implying that S is a function of q. This possibility has been confirmed by Hanushek, Lavy, and Hitomi (2008), who find dropout rates directly related to the value added of Egyptian schools. Our concern here, however, is largely how school quality changes over time. After we control for family factors and selectivity/ability, it seems doubtful that aggregate changes in such dropout behavior over time (as opposed to cross-sectional variations at a point in time) will lead to substantial problems in our estimation.

[14] As we discuss below, these trends have been much stronger for other countries compared to the United States. This fact shows up in the regression estimates.

promoted to greater schooling levels over time. Our time-specific measure of school attainment may capture not only the effects of schooling itself but also the decrease in school selectivity over time. If so, the contributions of more recent cohorts' schooling will be underestimated. We deal with this problem by directly including a measure of school selectivity across cohorts in each country.

Second, skill accumulation during school years is also affected by family inputs (e.g., Coleman et al. 1966). If family inputs have increased over time, then the contribution of more recent cohorts' schooling will be overestimated. We deal with this problem by including in X two sets of variables: indicators for mother's education attainment and the infant mortality rate (IMR) for the birth country at the time of birth. The IMR captures the average health condition during an individual's early childhood development, which in turn has a long-term impact on learning.

Third, individuals may gain or lose skills as a result of the aging process itself. If individuals tend to lose skills because of aging, then the contributions of earlier cohorts' schooling will be underestimated, and vice versa. We include in X a polynomial of age, which is not country specific, to control for this problem. This specification captures the idea that losing or gaining literacy skills due to physical and mental depreciation is a universal process.[15]

Finally, differential learning-by-doing at the workplace across countries could enter. The questions in the literacy tests of the survey, however, concern tasks of day-to-day life and are not job specific. Therefore, we assume that work experience has a limited role in affecting the performance in the tests and that omitting workplace learning does not bias the estimates of the contributions of schooling to the literacy skill.

Because primary-secondary schooling quality and college quality may vary over time in different manners, we also estimate equation (2) splitting the schooling variable into two parts: years of schooling before completing high school (≤ 12 years) and years of schooling after completing high school. These two variables are again country specific and cohort specific.[16]

We obtain a semiparametric estimate of q_{kc} by regressing the literacy score of each individual on observed years of schooling and the elements of X while allowing the coefficient on schooling to vary freely by country and cohort. Specifically, we estimate a modified version of equation (2):

$$L_{ikc} = \sum_{a=1}^{5} \beta_{ka} S_{ika} + X_{ik}\gamma_k + \varepsilon_{ikc}, \tag{4}$$

[15] See Smith and Marsiske (1997). The skill depreciation with the aging process could, of course, be distorted by different time patterns of nutrition and health care across countries, but we directly control for IMRs in part as a general measure of health environment.

[16] This specification is complementary to the earnings determination analysis by Heckman, Lochner, and Todd (2008) that suggests nonlinearities of postsecondary schooling in wage equations.

where $S_{ika} = 0$ if $a \neq c$. This estimation provides an estimate of how observed school attainment (S) translates into literacy scores for individuals in different age cohorts. From the five separate estimates of the schooling parameter in each country, we use the cohort aged 26–35 as the base group ($c = 2$) and estimate q_{kc} by equation (5):

$$\hat{q}_{kc} = \frac{\beta_{kc}}{\beta_{k2}}. \tag{5}$$

An index greater than one would indicate that cohort c's schooling is of higher quality than that of the base cohort; therefore, each year of schooling of cohort c would be equivalent to more than 1 year of the base cohort's schooling. With the estimate of school quality for each cohort, it is possible to estimate quality-adjusted schooling according to equation (3).

B. *Measuring School Selectivity*

Quality indices of schooling are derived from the estimated contributions of schooling at different time periods to the literacy skills. Understanding the pattern of school selectivity across countries, because it indicates varying ability of people with similar schooling at different times, is a first step.

We assume that in any given country, an individual of cohort c who completes school level s on average has higher ability than any individual of the same cohort who completes a school level less than s. If share ω of the population of cohort c completes at least school level s, then an individual of cohort c who completes school level s will have higher expected ability than share $1 - \omega$ of the population of the same cohort who stopped schooling earlier. We therefore assign $1 - \omega$ as the selectivity measure for an individual i of cohort c who completes school level s. We interpret this as effectively a broad index of ability for individual i.

The intuition behind this approach is straightforward. Assume, say, that there was no higher education in a country in 1950 but that by 1990 one-third of students continued studies into college. In 1950, secondary school graduates would include all of the country's most able individuals, but if college positions were allocated by ability as with national test scores, none of the most able would be in the group that stopped schooling at the secondary level in 1990. Indeed, the results can be seen from tracing out over time average scores on the Scholastic Aptitude Test in the United States. As the proportion of the population voluntarily taking the SAT rose in the 1960–80 time period, mirroring the increased attendance in higher education, the test became less selective and the average scores were depressed (Congressional Budget Office 1987).

For the United States in 1994, 86 percent of individuals between ages 26 and 35 completed at least high school education, and 14 percent did not finish high school. Therefore, if individuals are sorted by ability, somebody of this cohort who completed exactly high school would be expected to have ability higher on average than 14 percent of the cohort and is assigned a selectivity index of 0.14. Similarly, the selectivity index of an average American in this cohort that completed at least college education is 0.68.

OECD (2005) provides historical information about completion of upper secondary and tertiary schooling by different age groups across countries.[17] These data permit us to calculate selectivity indices across countries for individuals of different ages. The data on completing tertiary education for some countries and some cohorts are also divided between academic and vocational-technical.[18] The selectivity measures can range between zero and one, with individuals completing less than upper secondary education receiving a selectivity measure of zero.

Across sampled countries and over time, the selectivity of schooling shows wide variation. Tables 2 and 3 provide the selectivity measures $(1 - \omega)$ for five cohorts of individuals completing upper secondary education and completing tertiary education for each country. Table 2 aggregates all tertiary schooling, and table 3 separates vocational-technical from academic where available. While the United States has seen little change in the selectivity of schooling over the 5 decades represented in table 2, other countries, such as Poland and Sweden, have had dramatic changes. The strong trend toward more schooling across most of the countries implies that individuals at each level of schooling from earlier cohorts have higher selectivity measures (higher average ability) than those from more recent cohorts. This changing selectivity is most pronounced for individuals completing upper secondary education, as countries have expanded secondary education at a much faster pace than tertiary education.

Note again the purpose. We wish to estimate cohort-specific quality measures. In doing this, we are most worried about other possible explanations of achievement that have varied across time and that might

[17] Historical data come from a variety of OECD publications including various years of *Education at a Glance* (e.g., OECD 2005) and OECD (1995). Because we use the population survey information in one year to characterize populations going to school at different time periods, we need to assume that the underlying distribution for each cohort is stable over time. This assumption may not hold if, e.g., the proportion of immigrants in a cohort has changed considerably over time.

[18] See OECD (2005, annex 3) for a description of the International Standard Classification of Education. In short, type B tertiary education is generally practical-technical-occupational oriented with a minimum duration of 2 years and does not prepare students for more advanced study (vocational-technical). Type A tertiary education is more theoretically oriented with a minimum duration of 3 years and is intended to provide sufficient qualifications for gaining entry into advanced research programs and professions with high skills requirements (academic).

TABLE 2
SELECTIVITY MEASURES $(1 - \omega)$ FOR 10-YEAR COHORTS WITH DIFFERENT SCHOOLING LEVELS: TERTIARY SCHOOLING AGGREGATED (Most Selective = 1)

	Completed Upper Secondary					Completed 3 or More Years Tertiary				
Country	16–25	26–35	36–45	46–55	56–65	16–25	26–35	36–45	46–55	56–65
Canada	.10	.18	.21	.30	.47	.47	.49	.51	.55	.69
Chile	.36	.45	.55	.65	.76	.82	.89	.90	.91	.95
Czech Republic	.06	.08	.12	.16	.24	.86	.90	.88	.90	.92
Germany	.15	.10	.12	.16	.28	.78	.80	.73	.76	.83
Hungary	.15	.23	.27	.35	.69	.80	.86	.86	.86	.90
Ireland	.22	.39	.53	.65	.73	.63	.76	.81	.85	.89
Italy	.34	.45	.50	.65	.81	.84	.91	.89	.91	.95
Netherlands	.24	.31	.36	.46	.56	.73	.76	.75	.81	.86
Norway	.06	.06	.11	.21	.32	.59	.67	.71	.74	.79
Poland	.12	.12	.18	.32	.53	.90	.90	.90	.88	.92
Sweden	.09	.15	.22	.31	.48	.59	.73	.70	.74	.83
Switzerland (French, German)	.12	.11	.16	.21	.27	.69	.78	.77	.78	.83
United States	.14	.14	.11	.15	.24	.62	.68	.64	.67	.76

Source.—Author calculations from OECD (2005).
Note.—Bold indicates countries used in subsequent earnings analysis.

distort our estimates of school quality.[19] For these purposes, imperfect selection into higher education is less a concern than aggregate changes over time.

C. *International Patterns of School Quality Changes*

The school quality regression is based on individual observations for the average literacy scores in prose, documentary, and quantitative skills (normalized to have mean zero and standard deviation one within each country to facilitate interpretation). The sample includes individuals between 16 and 65 years of age who either are born in the survey country or immigrate to the survey country with at most primary schooling before immigration. The explanatory variables of primary interest are the country- and cohort-specific years of schooling for five 10-year age cohorts: 16–25, 26–35, 36–45, 46–55, and 56–65. We control for gender, selectivity of schooling, IMR, mother's education,[20] age, and age

[19] A related issue is whether aggregate abilities have changed across time. Specifically, Flynn (1984, 1987) and others suggest that IQ scores have risen over the time period covered by this analysis. The rise in measured IQ scores is now frequently referred to as the "Flynn effect." The important issue here is the strength of such movement that is independent of mother's education and health issues that are included here. We return to this issue below in the empirical discussion.

[20] Mother's education is measured by seven indicators for no education or some preprimary education, completing at most primary education, completing at most lower secondary education, completing at most upper secondary education, some college, completing college and above, and information not available.

TABLE 3

Selectivity Measures $(1 - \omega)$ for 10-Year Cohorts with Different Schooling Levels: Tertiary Schooling Disaggregated

(Most Selective = 1)

	Completed Upper Secondary					Completed Vocational-Technical Tertiary					Completed Academic Tertiary				
Country	16–25	26–35	36–45	46–55	56–65	16–25	26–35	36–45	46–55	56–65	16–25	26–35	36–45	46–55	56–65
Belgium	.19	.30	.42	.53	.69	.60	.67	.73	.78	.87	.81	.86	.89	.90	.94
Denmark	.13	.15	.20	.22	.33	.60	.73	.73	.73	.81	.69	.93	.95	.95	.97
Finland	.11	.15	.20	.36	.50	.62	.63	.66	.72	.81	.73	.85	.85	.87	.92
Great Britain	.14	.14	.20	.28	.41	.77	.77	.76	.79	.84	.85	.85	.85	.88	.92
New Zealand	.15	.36	.36	.45	.53	.69	.76	.72	.74	.79	.74	.86	.87	.90	.94
Switzerland (Italian)	.12	.12	.17	.20	.29	.69	.75	.75	.78	.82	.78	.84	.85	.87	.89

Source.—Author calculations from OECD (2005).

Note.—Bold indicates countries used in subsequent earnings analysis.

squared. Because we control for selectivity (interpreted as aggregate expected ability differences), family inputs, and the aging process, the coefficient estimate on a cohort-specific schooling measure is interpreted as the contribution to cognitive skills of an extra year of schooling of the cohort, which provides the foundation for estimating the quality of schooling received by the cohort from equation (5) and for estimating quality-adjusted schooling from equation (3).

Over time, female educational attainment has improved dramatically: 2.2 percent of the mothers of the oldest cohort had some tertiary education and 34.5 percent secondary education, whereas 9.5 percent of mothers of the youngest cohort had some tertiary education and 55.5 percent secondary education. Along with the increase of mother's education attainment, we expect learning environment within families to improve and children to obtain more human capital ceteris paribus. The IMR variable is obtained from Tamura (2006). We use the IMR measure at 5-year intervals; the IMR of 1930 is for the oldest individuals in the sample born between 1929 and 1934 and the IMR of 1980 is for the youngest individuals born after 1980. It has decreased for all the birth countries in the sample, and Appendix figure A1 depicts the pattern of decrease for the 13 countries that we focus on.[21]

Coefficient estimates on cohort-specific years of schooling, selectivity, and IMR for each country are reported in table 4. For example, for the cohort aged 26–35 in the United States, one more year of schooling increases one's cognitive skill by 0.134 standard deviations. Also reported is the p-value for the F-test that one more year of schooling has the same contribution to cognitive skills for each cohort. For most countries, schooling's contribution to cognitive skills has increased gradually over time; the increase from the cohort aged 56–65 to the cohort aged 16–25 ranges from 29 percent in Denmark and the Czech Republic to 220 percent in the Netherlands, and the trend is statistically significant. For Chile, Poland, the United States, and Italy, the upward trend in the point estimates for schooling's contribution over time is not present. In particular, in the United States, there is little difference in schooling's contribution across cohorts.

The coefficient estimates on cohort-specific years of schooling suggest that education quality has increased steadily over time for most of the European countries. A plausible, albeit speculative, explanation for this pattern of change in school quality relates to World War II and its aftermath. Countries experiencing significant quality improvement tend to be those deeply involved in World War II; their education system experienced severe damage and disruption during the war and had to be reconstructed in the postwar period. The oldest cohort, who received

[21] In the regression analysis, we also experimented with IMR obtained from the United Nations and per capita GDP from Maddison (1995) as a proxy for aggregate health condition in the birth country at the time of birth; the results are broadly similar. Further discussion of the IMR measure and its effect on the sample size is left to App. A.

TABLE 4

Contribution of School Attainment and Selectivity to Literacy Scores by 10-Year Age Cohorts

Age Cohort	Chile	Czech Republic	Denmark	Finland	Germany	Hungary	Italy	Netherlands	Norway	Poland	Sweden	Switzerland	United States
16–25	.169	.165	.153	.119	.073	.159	.142	.067	.191	.136	.092	.103	.130
	[.008]**	[.014]**	[.010]**	[.008]**	[.012]**	[.013]**	[.011]**	[.008]**	[.013]**	[.012]**	[.010]**	[.009]**	[.011]**
26–35	.159	.146	.131	.094	.071	.138	.143	.055	.177	.125	.089	.090	.134
	[.007]**	[.013]**	[.008]**	[.006]**	[.010]**	[.013]**	[.011]**	[.007]**	[.011]**	[.010]**	[.009]**	[.008]**	[.010]**
36–45	.154	.131	.129	.090	.060	.125	.145	.047	.174	.134	.069	.078	.138
	[.007]**	[.012]**	[.008]**	[.006]**	[.011]**	[.012]**	[.011]**	[.007]**	[.011]**	[.009]**	[.008]**	[.008]**	[.010]**
46–55	.162	.133	.118	.097	.056	.110	.155	.042	.168	.135	.069	.079	.142
	[.008]**	[.013]**	[.008]**	[.007]**	[.011]**	[.011]**	[.012]**	[.008]**	[.012]**	[.009]**	[.008]**	[.009]**	[.009]**
56–65	.162	.128	.119	.089	.034	.090	.148	.021	.141	.137	.053	.074	.143
	[.009]**	[.015]**	[.010]**	[.009]**	[.014]*	[.013]**	[.013]**	[.009]*	[.013]**	[.010]**	[.010]**	[.011]**	[.010]**
p-value	.04	0	0	0	.11	0	.27	0	0	.07	0	0	.49
Selectivity	−.050	.172	.301	.586	.630	.226	−.237	.891	−.007	.161	.435	.242	.112
	[.072]	[.094]+	[.066]**	[.068]**	[.107]**	[.108]**	[.111]*	[.069]**	[.099]	[.072]*	[.078]**	[.065]**	[.093]
IMR	−2.146	−3.314	−9.473	−13.794	−7.402	−1.173	−8.029	−10.318	−10.927	−5.209	−7.418	−10.706	−10.463
	[.708]**	[1.301]*	[2.181]**	[2.050]**	[3.007]*	[1.447]	[2.053]**	[1.481]**	[2.792]**	[1.131]**	[3.816]+	[2.586]**	[1.811]**

Note.—Robust standard errors are in brackets. The sample includes individuals between 16 and 65 years of age, native born or immigrants completing at most primary education before immigration. The dependent variable is the normalized average literacy skill test score. Control variables are age, age squared, country-specific school selectivity, country-specific infant mortality rate (IMR), country-specific indicators for mother's education level, and a country-specific indicator for female. Education's contributions to literacy skills for different cohorts are the coefficient estimates on the interactive terms between education (measure by total years of schooling) and indicators for the respective age cohorts. *p*-value is for the *F*-test that education's contributions to literacy score are the same over the five cohorts.

+ Significant at 10 percent.

* Significant at 5 percent.

** Significant at 1 percent.

their education during or immediately after the war, would have suffered the most. This disruption, for example, seems to drive the extraordinarily large quality improvement for the Netherlands, where the most significant change comes from the 55–65-year-olds (who would have been born in 1930–40). With the national education system gradually back to normal, the quality of education increased for the subsequent cohorts. The education systems in the United States, Chile, and Italy were relatively undisrupted during the war, perhaps supporting the stable quality of schooling over this period.[22]

Individuals with more educated mothers tend to have higher cognitive skills (not separately shown).[23] Higher IMR consistently predicts lower cognitive skills, and the relationship is significant for all countries but Hungary. Selectivity is also important in tracking cognitive skills across cohorts. Our measure of selectivity of different school attainment has a positive effect on cognitive skills for 10 out of the 13 countries and is statistically significant at the 5 percent level for all but the United States and the Czech Republic.

The coefficients on age and age squared (common to all countries) are 0.013 and −0.00053, respectively; they are jointly significantly different from zero. By these estimates, cognitive skills increase with age for all individuals in our sample (between 16 and 65 years of age). This pattern is consistent with findings in the literature on psychology of adult learning (see Smith and Marsiske [1997] and references therein).

One concern is that the quality of primary-secondary schooling and the quality of college education evolve differently. We address this concern by splitting years of schooling into two parts: years of primary-secondary schooling ($S \leq 12$ years) and years of tertiary education ($S > 12$ years). Appendix table B1 reports the coefficient estimates on cohort-specific primary-secondary schooling and tertiary schooling for each country, again providing estimates of the contribution of one extra year of primary-secondary schooling or college education to the cognitive skills for different cohorts. For virtually every cohort in every country, primary-secondary schooling has a much bigger contribution to cognitive skills than college education.[24] This is expected given that the skills tested by the IALS are day-to-day tasks and are more directly affected by basic education. For the same reason, the trends observed

[22] See U.S. Office of Education (1945) and Lowe (1992). While Italy was clearly a combatant, Italy's wartime experience apparently had minimal effect on the schools, and the postwar reconstruction proceeded rapidly; see Wolff (1992).

Poland is an exceptional case. As is clear in the second-stage wage regression, the cognitive skill measure never plays a significant role in explaining wage earnings, in contrast to all the other countries. This prompts us to suspect that there is some serious measurement issue related to the cognitive skill measure for Poland.

[23] The impact of mother's education is estimated semiparametrically for each country according to major division of schooling (e.g., primary or lower secondary). The precise division of mother's education depends on the detailed survey data for each country.

[24] Only the youngest cohort in Germany and Sweden is the exception.

in table 4 reflect to a large extent the evolution of the primary-secondary schooling quality, as displayed in Appendix table B1. The estimated changes in quality of tertiary schooling are insignificant for half the countries in the sample, although this could simply reflect the much smaller samples of tertiary graduates than those of primary and secondary schooling.

As discussed earlier, the major challenge in estimating the school quality changes over time from equation (2) is to adequately control for other factors that may be correlated with both schooling and literacy over time and across cohorts. In obtaining results in table 4, we have controlled for major influences including IMR at birth, mother's education, and selectivity of school system; however, one potentially important factor we have not controlled for is IQ. In a series of influential papers, James Flynn (1984, 1987) found that IQ had increased by about 2–3 points per decade in various developed countries since the 1930s—a phenomenon that has been dubbed the "Flynn effect." If IQ is positively correlated with both schooling and literacy over time, then not controlling for it may lead to an overestimate of school's contribution to literacy.

Directly controlling for IQ for different age groups in the regression turns out to be quite difficult. Historically, IQ tests were most commonly administered in the United States and Great Britain, where representative samples of the population were tested at various times (Flynn 1984, 1987). IQ tests in other countries were less frequent and were usually given to a specific subpopulation or very small samples. For example, the "strong" data reported in Flynn (1987) for the Netherlands and Norway (both in our sample) were based on tests administered by the military to 18- or 19-year-old males, whereas the "weak" data for Germany and Switzerland were based on tests given to small samples of children of varying age ranges and were normalized under strong assumptions. Lynn and Vanhanen (2002) created average IQ measures for 185 countries; however, for the 81 countries for which they actually have historical IQ data (including Italy, Denmark, Finland, and Sweden in our sample), the IQs were observed at just two different points in time, and they were from small samples and for different age groups. Using these existing measures to create an IQ variable for each age by extrapolation requires strong assumptions about the evolution of IQ for different age groups and for most countries will lead to a variable that is highly correlated with age.

Because of this difficulty in creating an IQ variable for all the countries in our sample, we chose instead to explore the impact of IQ focusing on the United States. IQ measures for the United States compiled by Flynn (1984) were based on large samples representative of the U.S. population and had relatively many observations. From these measures, we interpolated IQs for other ages. We estimate our intertemporal literacy equations for the United States both with and without controlling for IQ and obtain similar estimates on time-specific years of schooling; indeed, the two sets of estimates are not statistically significantly different from each other. Additionally, the estimate on IQ is insignificant.[25] Thus, controlling for IQ does not appear to affect the estimates of U.S. schooling quality.

Furthermore, recent research suggests that measured IQ is strongly influenced by environmental factors such as socioeconomic status, mother's education, and nutrition and by the aging process (Neisser 1998; Flynn 2007).[26] We already control for mother's education, IMR, and age, which likely explains the insignificance of IQ in our U.S. estimates and leads us to be less concerned about potential bias in our literacy estimates for other countries.

Taking the cohort aged 26–35 as the base cohort, we construct a quality-adjusted measure of years of schooling as defined in equation (3), using the cohort-specific estimate of education's contribution to literacy skills reported in table 4. Our quality-adjusted schooling measure is used to determine the lifetime earnings gradients to different levels of schooling for the base cohort.[27]

V. Skill Gradients in the Labor Market

We now turn to the estimation of quality-consistent labor market impacts of schooling and literacy skills. We first apply a standard Mincer framework using the quality-adjusted years of schooling measure to proxy for individual skills as in equation (6):

$$\ln (y_{ikc}) = \delta_k \cdot \tilde{S}_{ikc} + Z_{ikc} \cdot \theta_k + v_{ikc}. \tag{6}$$

The dependent variable, $\ln (y_{ikc})$, is the logarithm of annual earnings from employment in the survey year of individual i; $\tilde{S}_{ikc}$ is individual i's quality-adjusted years of schooling; Z_{ikc} is a vector of control variables, including an indicator for female, potential experience, and an indicator for living in a rural area; and θ_k is the vector of relevant country-specific parameters. Because schooling is normalized relative to the quality of schooling of the cohort aged 26–35, the coefficient estimate of $\delta_k = \partial \ln y/\partial \tilde{S}$ measures the lifetime proportionate increase in earnings expected for an additional year of schooling for this base cohort in

[25] Regression results here are not completely comparable to the regression of eq. (2), in part because eq. (1) assumes the same aging process for all countries, whereas here we assume that the aging process is country specific; this assumption aggravates the problem of collinearity between IQ and age. The collinearity problem, however, does not affect estimates on time-specific years of schooling. Appendix table B2 shows regression results using U.S. data only (with and without IQ) and comparable results for the United States from table 4.

[26] See also the environmental models of Dickens and Flynn (2001).

[27] A quality-adjusted schooling measure constructed from App. table B1 that allows for heterogeneous effects of higher education attainment is very similar, given the closeness of two different estimates.

country k. While $\tilde{S}_{ikc}$ is our preferred schooling measure, we provide separate estimates of equation (6) with and without quality adjustments to schooling. In this way, we can relate our estimates to the common alternatives in the literature.

This earnings gradient, which we often describe simply as the return to schooling, reflects how skills are valued in the different countries. It is not, however, a direct estimate of the internal rate of return (IRR) to an added year of schooling. As Heckman, Lochner, and Todd (2008) point out, proper calculations of an IRR for individuals would take into account direct costs and taxes and, by our analysis, the differential opportunity costs for some cohorts of obtaining a quality-consistent year of schooling.

In the earnings analysis, we focus on the 13 countries with continuous wage measures in the IALS.[28] We estimate the returns to education using the sample of individuals working full-time during the 12 months prior to the survey. Full-time workers are defined as those working at least 40 weeks and at least 30 hours per week during the previous 12 months.[29] Whenever we include quality-adjusted schooling, we rely on the estimates from the full IALS sample that were reported in table 4.

A. *Wage Gradients for Quality-Adjusted Schooling*

As a benchmark, we first estimate a classical Mincer wage equation within each country using actual years of schooling as the measure of the quantity of human capital, controlling for gender, potential experience and its square, and an indicator for living in a rural area. As noted, we follow convention by referring to the schooling gradient as the "return to schooling," with the caveat that this does not imply acceptance of the Mincer simplifying assumptions.

The return to schooling for each country is reported under model 1 in table 5. One extra year of education increases annual earning by from 3.8 percent in Sweden to 11.3 percent in the United States with an unweighted average across all countries of 6.5 percent. Educational attainment is considerably more highly rewarded in the United States than in other developed countries, consistent with findings in the literature. Also noticeable is that the return to schooling in the four countries at the bottom of development in the sample (Poland, Czech Re-

[28] Three countries (regions) with continuous wage measures are not included in the wage analysis for different reasons. Canada does not have an age measure, Slovenia does not have historical information on schooling patterns for estimation of the selectivity measure, and Italian-speaking Switzerland has too few observations.

[29] For Sweden, the full-time working status is based on answers to questions of whether a respondent works and whether he or she works full-time.

TABLE 5
ALTERNATIVE ESTIMATES OF THE RETURNS TO SCHOOLING AND LITERACY SKILLS IN THE LABOR MARKET

	Chile	Czech Republic	Denmark	Finland	Germany	Hungary	Italy	Netherlands	Norway	Poland	Sweden	Switzerland	United States
Model 1:													
Schooling	.110	.062	.056	.048	.049	.077	.055	.051	.053	.083	.038	.054	.113
	[.010]**	[.007]**	[.004]**	[.006]**	[.009]**	[.009]**	[.006]**	[.004]**	[.006]**	[.008]**	[.006]**	[.006]**	[.009]**
Model 2:													
Quality-adjusted schooling	.109	.074	.060	.045	.066	.093	.052	.065	.057	.074	.063	.065	.105
	[.010]**	[.008]**	[.005]**	[.006]**	[.011]**	[.011]**	[.006]**	[.006]**	[.007]**	[.008]**	[.009]**	[.008]**	[.008]**
Model 3:													
Quality-adjusted schooling	.088	.065	.050	.036	.057	.082	.044	.047	.049	.072	.058	.048	.074
	[.011]**	[.009]**	[.006]**	[.006]**	[.011]**	[.012]**	[.006]**	[.006]**	[.007]**	[.008]**	[.010]**	[.009]**	[.010]**
Literacy	.134	.050	.066	.094	.079	.072	.067	.172	.059	.013	.038	.153	.197
	[.033]**	[.017]**	[.015]**	[.024]**	[.020]**	[.032]*	[.021]**	[.020]**	[.020]**	[.025]	[.022]$^{+}$	[.029]**	[.037]**
Model 4:													
Schooling	.089	.054	.047	.039	.042	.069	.047	.038	.045	.082	.034	.040	.080
	[.011]**	[.008]**	[.005]**	[.006]**	[.009]**	[.009]**	[.007]**	[.004]**	[.006]**	[.009]**	[.006]**	[.007]**	[.011]**

Literacy	.131	.052	.064	.086	.082	.065	.066	.158	.056	.008	.049	.152	.193
	[.033]**	[.017]**	[.015]**	[.024]**	[.020]**	[.032]*	[.022]**	[.021]**	[.020]**	[.025]	[.022]*	[.029]**	[.038]**
Model 5:													
Quality-adjusted schooling	.070	.044	.039	.003	.059	.039	.046	.028	.036	.061	.038	.044	.062
	[.019]**	[.016]**	[.007]**	[.010]	[.014]**	[.020]+	[.019]*	[.007]**	[.014]*	[.018]**	[.012]**	[.011]**	[.012]**
Ability	.054	.124	.167	.350	.324	.502	−.019	.536	.241	.198	.236	.313	.464
	[.150]	[.105]	[.042]**	[.080]**	[.102]**	[.147]**	[.140]	[.056]**	[.104]*	[.113]+	[.078]**	[.073]**	[.126]**
Observations	1,375	1,361	1,688	1,342	680	896	954	1,190	1,487	1,154	1,343	998	1,158

Note.—Robust standard errors are in brackets. The sample includes full-time workers between 16 and 65 years of age, native born or immigrants completing at most primary education before immigration. In all the models, the dependent variable is the logarithm of annual earnings from employment; control variables are gender, potential experience and its square, and an indicator for living in rural area. In model 1, education is measured by actual years of schooling. In model 2, education is measured by quality-adjusted years of schooling, where the quality index of schooling is derived from education's contribution literacy skills. Model 3 is model 2 controlling for individual literacy skill. Model 4 is model 1 controlling for individual literacy skill. In model 5, education is measured by quality-adjusted years of schooling, and other controls include ability, IMR of birth county at birth, and indicators for mother's education but does not include literacy scores.

+ Significant at 10 percent.

* Significant at 5 percent.

** Significant at 1 percent.

public, Hungary, and Chile) is much higher than that in the more developed European countries.[30]

The classical Mincer framework makes use of variation in years of schooling received at different time periods, and, as demonstrated, the quality of the schooling is not comparable over time, making the Mincer estimates an average of the returns to education of different qualities. Model 2 of table 5 reports the estimated return to quality-adjusted schooling for the base cohort, the cohort aged 26–35 in the survey year for each country.[31]

While adjusting for secular changes in school quality makes little to no difference in Chile, Italy, and the United States, it substantially alters the estimated earnings growth with schooling in the remaining countries. The most salient difference between models 1 and 2 in table 5 is that there is a significant increase in the return to education for a majority of countries. The increase in returns is over 30 percent in Germany and Sweden, and Hungary, Switzerland, and the Netherlands also show an increase in excess of 20 percent of the standard Mincer estimate. Accounting for quality movements on average increases the estimate of the earnings gradient by 10 percent from the basic Mincer return.[32]

Estimates of returns to quality-adjusted schooling in model 2 have smaller variation across countries than those for measured schooling in model 1. In particular, the gap in the return to education between the United States and other countries becomes smaller once the change in education quality is taken into account. This is readily seen in figure 2, which plots the unadjusted and adjusted Mincer returns across the 13 countries. This convergence of estimates suggests that the much higher reward to education in the United States relative to other countries is in part an artifact of the stable quality of its education system. With large improvement in the education system of other countries, the gap in the return to education is noticeably smaller for today's graduates.

With the exception of the United States, the high returns to schooling are systematically found in the less developed countries in our sample. The countries with more developed welfare states fall in the lower range of returns, but this is not just due to higher taxes because these results are all pretax earnings.

[30] In 1995, Poland, Hungary, and Chile each had GDP per capita below the world average. The Czech Republic was slightly above this average. Each was roughly one-quarter or less of the average in the first 12 E.U. countries. See international data from the Economic Research Service of the U.S. Department of Agriculture (http://www.ers.usda.gov/Data/Macroeconomics/).

[31] Standard errors for models 2 and 3 are robust standard errors. Bootstrap standard errors show little change, suggesting that the problem of standard error estimation with derived explanatory variables discussed in Murphy and Topel (1985) is negligible in the present context.

[32] The difference in returns to education (model 1) and quality-adjusted education (model 2) is significant at the 5 percent level for all countries but Finland and Chile.

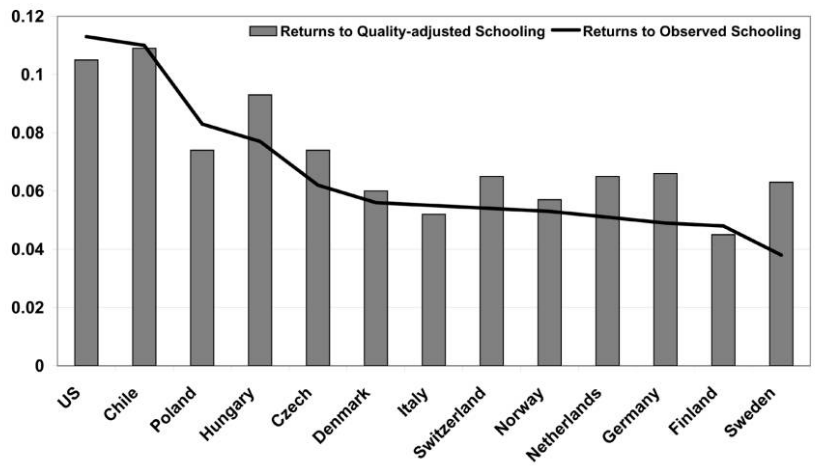

Figure 2.—Impact of school quality adjustment on estimated labor market returns to schooling.

B. *Returns to Cognitive Skills*

The above adjustment for school quality makes schooling comparable over time, but this schooling measure itself does not provide a complete measure of individual skills. Individuals with a given number of quality-adjusted years of schooling will still show very different cognitive skills at any given time, both because of cross-sectional variation in school quality and because of other factors affecting achievement. Adding a measure of cognitive skills in the Mincer earnings function permits direct investigation of how the labor market values cognitive skills, which is produced by schooling and other inputs. We thus estimate equation (7):

$$\ln (y_{ikc}) = \delta_{1k} \cdot \tilde{S}_{ikc} + \delta_{2k} \cdot L_{ikc} + Z_{ikc} \cdot \theta_k + v_{ikc}, \tag{7}$$

where L_{ikc} is individual i's normalized literacy test score. In this model, the literacy test score is intended to proxy for individual cognitive skills, whereas the quality-adjusted schooling $\tilde{S}_{ikc}$ measures the human capital differences that are not captured by L_{ikc}, such as noncognitive skills or cognitive skills that are imperfectly measured by L. From equation (2), L_{ikc} is determined by years of schooling and many other factors such as family inputs and ability. The return to measured cognitive skills is δ_{2k} in equation (7).

As shown in model 3 of table 5, the impact of cognitive skills is positive and statistically significant for all countries but Poland. A one-standard-deviation increase in the literacy score increases annual earnings by from 3.8 percent in Sweden to 20 percent in the United States. Figure 3 shows the returns to literacy scores across countries. Contrary to the

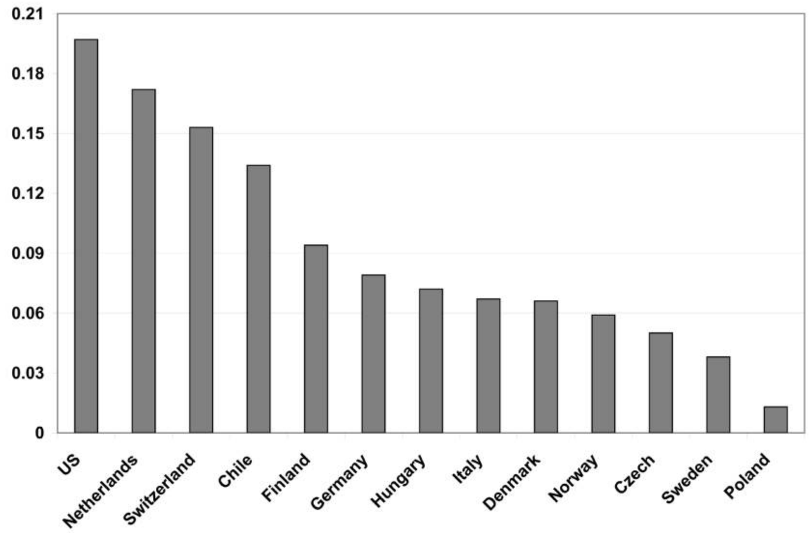

Figure 3.—Estimated labor market returns to cognitive skills

pattern of returns to school attainment in figure 2, there is no obvious pattern by stage of development as to where the returns to cognitive skills are high or low.[33]

The estimates of the returns to skill in the IALS data for the United States are significantly above those in recent studies. When separate panel data on returns to cognitive skills early in a career are used, three different estimates point very consistently to a return of about 12 percent per standard deviation (see Mulligan 1999; Murnane et al. 2000; Lazear 2003).[34] Our larger estimates for the United States may reflect the returns that accrue later in the working life and that are not observed in these panel data estimates.[35]

[33] An obvious follow-on research project is investigating what aspects of these economies lie behind the pattern of returns to cognitive skills.

[34] Murnane et al. (2000) provide evidence from the High School and Beyond and the National Longitudinal Survey of the High School Class of 1972. Their estimates suggest some variation, with males obtaining a 15 percent increase and females a 10 percent increase per standard deviation of test performance. Lazear (2003), relying on a somewhat younger sample from the National Education Longitudinal Study of 1988, provides a single estimate of 12 percent. These estimates are also very close to those by Mulligan (1999), who finds 11 percent for the normalized Armed Forces Qualifying Test score in the National Longitudinal Survey of Youth data.

[35] Altonji and Pierret (2001) suggest that the role of cognitive skills—which are difficult for an employer to observe—may grow with the worker's experience in the labor market. At initial hiring, the employer relies more on the observable measures of school attainment, but as time goes on, the employer can substitute direct observations of worker skills (measured here by literacy scores) for the cruder proxy of years of schooling. When we investigate this model within our data by permitting the impact of literacy scores to grow with age, we find some support for this statistical discrimination model for the United States, which is consistent with the different results here and in the previous panel studies.

After we control for cognitive skills, the coefficient estimate on years of schooling drops in all countries. This estimate, however, does not have an interpretation of the return to schooling: part of the return is reflected by the return to cognitive skills.[36] It is not comparable to the return to schooling reported in model 2 of table 5.[37] The estimate can be interpreted as the return to other skills produced by schooling such as noncognitive skills, which is still positive and significant.

For comparison, we estimate the wage equation using the actual years of schooling measure along with literacy skills. The results—similar to a number of comparable U.S. studies—are reported in table 5 under model 4. Compared to model 1, the estimate on years of schooling drops significantly; the return to literacy skills, however, is almost identical to that in model 3. This suggests that direct measures of cognitive skill such as literacy scores provide a stable estimate of returns to skills, reinforcing the stability of U.S. estimates across studies.

C. *Adjustment for Other Skill-Generating Factors*

As indicated by equation (2), schooling is just one of the many factors in generating cognitive skills; similarly, other skills such as noncognitive skills are also produced jointly by schooling and other factors such as family input and ability. If we follow the tradition and interpret the coefficient estimate δ_k on $\tilde{S}_{ikc}$ in equation (5) as returns to one extra year of schooling, δ_k will reflect the contribution to human capital of both schooling and these other factors and will be an upwardly biased estimate of return to schooling in the labor market. This is the same concern underlying the large literature that has concentrated on dealing with the ability bias in estimating returns to schooling (e.g., Card 1999; Glewwe 2002). In that literature, factors such as family inputs and ability positively affect the years of schooling an individual will obtain; therefore, without controlling for these factors, the return to schooling will be biased upward.

To estimate a return to schooling purged of other factors, we estimate a wage equation augmented by vector X from equation (2):

$$\ln (y_{ikc}) = \delta_k \cdot \tilde{S}_{ikc} + X_{ikc} \cdot \varphi_k + Z_{ikc} \cdot \theta_k + v_{ikc}, \tag{8}$$

We do not, however, find any support for the statistical discrimination model in countries other than the United States.

[36] A second complication is the possibility that schooling affects noncognitive skills that are important for earnings but are uncorrelated with the measured cognitive skills. Hanushek and Woessmann (2008) discuss the more general interpretation of the estimated parameters of eq. (5) in the presence of such noncognitive skills. For a discussion of the possible measurement and importance of these, see also Bowles et al. (2001), Cunha et al. (2006), and Heckman, Stixrud, and Urzua (2006).

[37] A broadly comparable rate of return to schooling can be recursively calculated from the coefficient estimates in eqq. (5) and (1) as $\hat{\delta}_{1k} + \hat{\delta}_{2k} \cdot \hat{\beta}_k$, where the second term reflects the return to schooling through the return to cognitive skills.

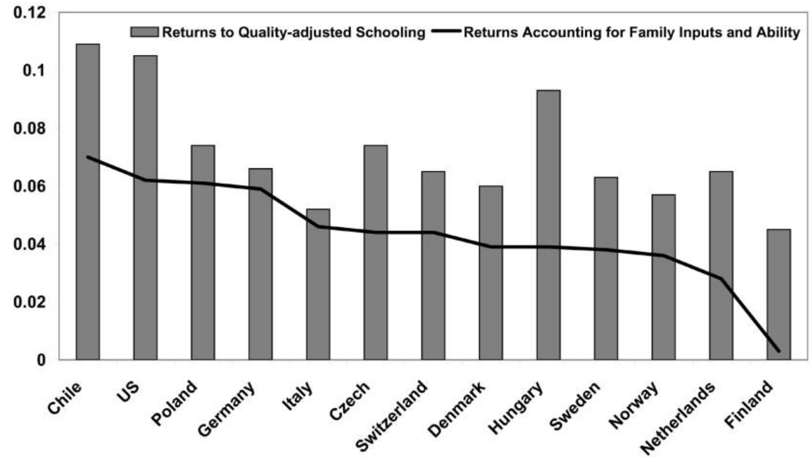

Figure 4.—Impact of controlling for family inputs and ability on estimated labor market returns to schooling.

where X_{ikc} includes the cohort ability measure, IMR at birth, and mother's education. The coefficient estimates on quality-consistent schooling and ability are reported under model 5 in table 5 for each country. To the extent that we have adequately controlled for the other influences on individual skills, $\hat{\delta}_k$ is now an estimate of the marginal impact of quality-adjusted years of schooling on earnings. Thus, the coefficient on the quality-consistent schooling is comparable to estimates in model 2 of table 5. In all countries, the return to schooling drops significantly after controlling for factors in X_{ikc}. This change is clearly illustrated in figure 4, where the solid line indicates the returns to schooling purged of the direct family and ability influences. The average return to one extra year of schooling over all countries is 4.8 percent. The coefficient on ability is positive and significant for all countries but Italy, the Czech Republic, and Chile. Estimates on IMR and mother's education (not reported) vary greatly across countries. Coefficient estimates of X_{ikc} are jointly significant in all countries but Norway.

These estimates also provide an explanation for prior discussions about the heterogeneity of returns to years of schooling across individuals. Our estimates explicitly recognize other sources of human capital that, if ignored, appear as individual variations in school attainment. If we have accurately incorporated the impacts of these other elements of skill differences, the estimated coefficient on quality-adjusted school attainment will be the average impact of schooling per se on earnings.

Table 6 summarizes the impact of the adjustment for school quality on the estimate of return to schooling. Column 1 compares the return to schooling and the return to quality-adjusted schooling. Column 2 compares the return to quality-adjusted schooling before and after re-

TABLE 6
Impact of School Quality Adjustment on Labor Market Returns

Country	% Change with School Quality Adjustment[a] (1)	% Change Controlling for Family Inputs and Ability[b] (2)
Sweden	65.79	−39.68
Germany	34.69	−10.61
Netherlands	27.45	−56.92
Hungary	20.78	−58.06
Switzerland	20.37	−32.31
Czech Republic	19.35	−40.54
Norway	7.55	−36.84
Denmark	7.14	−35.00
Chile	−.91	−35.78
Italy	−5.45	−11.54
Finland	−6.25	−93.33
United States	−7.08	−40.95
Poland	−10.84	−17.57
Summary statistics:		
Mean	13.28	−39.16
Standard deviation	21.56	21.87
Minimum	−10.84	−93.33
Maximum	65.79	−10.61

[a] Calculated from table 5 as (return to quality-adjusted education − return to education)/return to education × 100 in models 1 and 2 of the wage regression.
[b] Calculated from table 5 as (return to quality-adjusted education with controls for family inputs and ability − return to quality-adjusted education)/return to quality-adjusted education × 100 in models 2 and 5 of the wage regression.

moving the confounding other factors contributing to human capital production. The return to schooling exhibits significant changes in both comparisons, suggesting that the return to schooling estimate is likely to be subject to varying measurement of schooling and model specification. This is in sharp contrast to the estimate of return to cognitive skills, which is similar regardless how schooling is measured.[38]

VI. Conclusion

The widespread use of the Mincer earnings model to assess the returns to schooling around the world is testimony to its power to summarize important aspects of human capital investment. It has been broadly used

[38] Note that the percentage changes calculated in table 6 are not invariant to the choice of base cohorts for the school quality adjustment. A change of base cohort is a linear adjustment to $\tilde{S}$ so that the estimates for the effects of cognitive skills and influences of other factors on earnings are unaffected by any base change; but the percentage change from returns to unadjusted schooling (col. 1) is directly affected by a scale change and the change will differ by country. We think that normalizing by recent school quality is most natural for most current policy deliberations.

to analyze earnings and income distribution questions both within and across countries. The interpretation, nevertheless, depends on the deceptively simple empirical assumption that individuals used in comparisons of schooling and earnings are otherwise similar. This paper not only considers a series of key issues about the "otherwise similar" assumption but also extends the analysis to a larger international context.

Microdata from the International Adult Literacy Survey provide a unique opportunity to investigate international differences in the labor market returns to skill. The consistent measures of cognitive skills for workers of different ages within 13 countries permit direct analysis of how time-series and cross-sectional differences in schooling quality affect common approaches used in estimating rates of return to skills. These estimates are, as noted, a fundamental ingredient in calculating an internal rate of return on investments by individuals and societies in schooling and other skill-enhancing programs.

In the time-series dimension, the concern is that the quality of schooling may have changed over time within a country. If so, treating people with a given level of schooling obtained at different points in time can lead to bias in the estimated returns to schooling with the direction depending on the pattern of school quality change.

We construct an education quality index from the contribution of schooling during different periods to cognitive skills (after also correcting for the selectivity of schooling across time for each country, trends in infant mortality rate for country of birth, and mother's education attainment). When we estimate wage equations using the quality-adjusted schooling measure, we find that the returns to schooling for current cohorts are noticeably higher than the return to the unadjusted schooling in most countries, with the earnings gradient for schooling underestimated by more than 30 percent in some countries. But quality adjustments have little effect on the United States, and once the quality trends are taken into account, the labor market impacts of schooling in other countries appear closer to those in the United States.

A more fundamental issue regarding the Mincer wage regression is that even individuals educated at the same period of time are likely to experience very different schooling quality and other inputs, especially in an international context. Therefore, schooling is far from perfect as a proxy for skills, and the estimated return to schooling does not reflect how labor markets reward skills. Cross-country comparisons based on returns to schooling are particularly misleading. We explore directly the return to cognitive skills. Cognitive skills play an important direct role in determining an individual's earnings. Returns to cognitive skills are positive and significant in all but one country and are the highest in the United States.

The other way to approach this is to investigate the returns to (quality-adjusted) schooling after allowing for other factors that enter into the determination of individual skills. Our estimates of the return to quality-adjusted schooling controlling for family and ability factors that contribute to skills formation are a generalization of attempts to deal with ability bias in estimating the return to schooling. Once confounding factors such as family inputs and ability are controlled for, the return to schooling drops considerably, suggesting the presence of substantial bias in naïve estimation of Mincer equations. The magnitude of such bias is estimated to be much larger than typically found in other estimations of the rates of return to schooling.

Finally, while this analysis has focused exclusively on measuring skills for the modeling of wage determination, it is equally clear that the issues come into play in a much wider range of studies. It is commonplace in analyses of other behavioral outcomes—be it such diverse things as voter participation, individual health outcomes, or migration behavior—to include measures of school attainment to control for human capital and other inputs. But these analyses are equally subject to concerns about the adequacy of measured school attainment, particularly when a causal interpretation is attached to the impact of schooling.

Appendix A

Data

The main regressions in this paper focus on individuals between 16 and 65 years of age in 13 countries. We further restrict our sample to include only individuals who either were born in the survey country or immigrated to the survey country with at most primary education before immigration. Seventy-two countries were identified as birth countries for individuals included in the sample; however, some immigrants did not specify their countries of birth. The sample size is 28,708.

The IMR variable is obtained from Tamura (2006). We use the IMR measure at 5-year intervals: IMR of 1930 is for the oldest individuals in the sample born between 1929 and 1934, IMR of 1935 is for individuals born between 1935 and 1939, and so forth. IMR of 1980 is for the youngest individuals born after 1980. For high-income countries and a few less developed countries (Chile, Hungary, and Czech Republic), IMR data are available for almost the entire period from 1930 to 1980, and limited interpolation is performed to create the entire time series. For other countries, IMR data are mostly available between 1950 and 1980. We create values for Poland for earlier years by a linear extrapolation. We also create values for individuals not stating their countries of birth by averaging available IMRs over all the countries; this, however, is done only for years between 1950 and 1980 because IMRs for earlier years are almost exclusively for high-income countries. This adjustment creates a sample of 28,477 individuals with values for all the variables used in the school quality regression, of which 473 are immigrants finishing at most primary education before they immigrated to the survey country.

IMR from the United Nations Population Division is available for all countries at 5-year intervals from 1955 on. Values for earlier years are created by a linear extrapolation. The time series thus obtained is rather different from that obtained from Tamura (2006); indeed, the latter shows a nonlinear trend in within-country changes over time. Appendix figure A1 depicts the IMR from Tamura (2006) for the 13 countries under study.

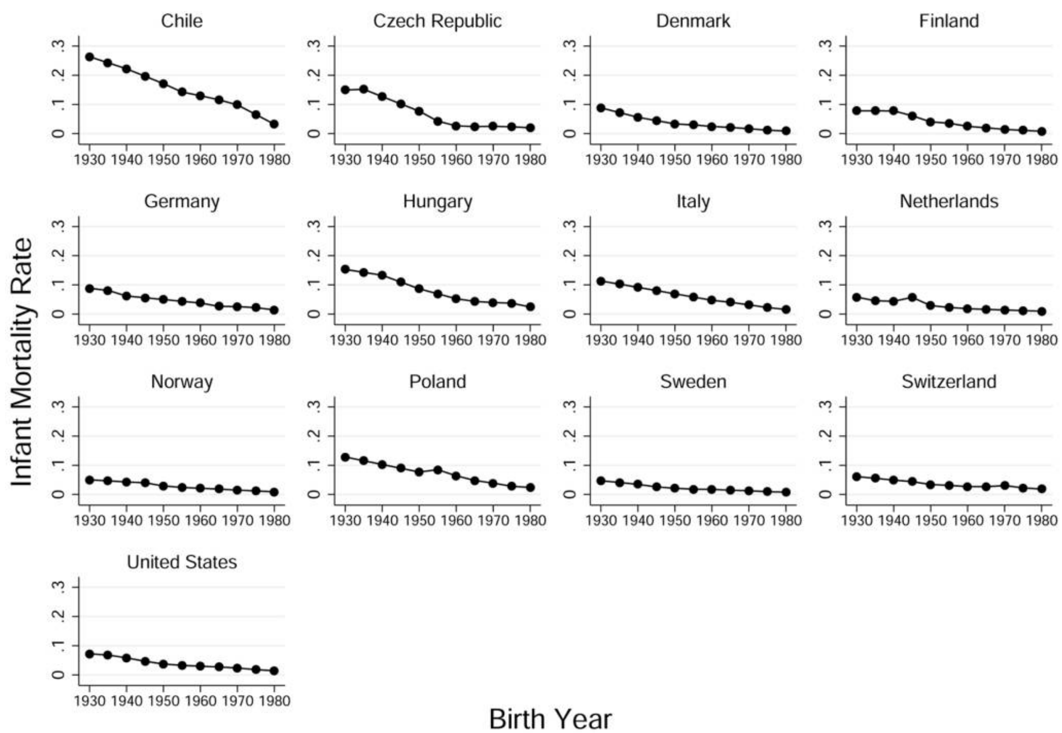

Figure A1.—Infant mortality rates by country and time

Appendix B

TABLE B1
Contribution of School Attainment Less than Tertiary and Tertiary to Literacy Scores by 10-Year Age Cohorts

Age Cohort	Chile	Czech Republic	Denmark	Finland	Germany	Hungary	Italy	Netherlands	Norway	Poland	Sweden	Switzerland	United States
							Less than Tertiary						
16–35	.170	.214	.235	.185	.068	.198	.217	.087	.226	.178	.077	.166	.191
	[.009]**	[.024]**	[.017]**	[.016]**	[.020]**	[.018]**	[.017]**	[.017]**	[.017]**	[.015]**	[.014]**	[.020]**	[.016]**
26–35	.160	.209	.216	.168	.084	.186	.228	.089	.225	.171	.084	.161	.205
	[.008]**	[.023]**	[.015]**	[.015]**	[.018]**	[.017]**	[.016]**	[.015]**	[.015]**	[.013]**	[.013]**	[.019]**	[.015]**
36–45	.162	.198	.216	.166	.076	.174	.227	.084	.228	.179	.073	.151	.214
	[.008]**	[.022]**	[.015]**	[.014]**	[.018]**	[.017]**	[.014]**	[.016]**	[.016]**	[.011]**	[.013]**	[.019]**	[.014]**
46–55	.181	.198	.195	.166	.076	.156	.228	.082	.222	.179	.081	.159	.231
	[.009]**	[.023]**	[.015]**	[.014]**	[.019]**	[.017]**	[.015]**	[.016]**	[.017]**	[.012]**	[.013]**	[.019]**	[.014]**
56–65	.178	.201	.183	.148	.055	.122	.210	.064	.188	.17	.059	.151	.230
	[.010]**	[.025]**	[.017]**	[.017]**	[.022]*	[.021]**	[.017]**	[.018]**	[.019]**	[.014]**	[.016]**	[.020]**	[.015]**
p-value (< tertiary)	.03	.40	0	.03	.11	0	.20	.28	0	.31	.03	.33	.01
							Tertiary						
16–25	.118	.155	−.003	.079	.133	.122	.083	.039	.140	.033	.122	.029	.107
	[.024]**	[.030]**	[.031]	[.018]**	[.029]**	[.030]**	[.018]**	[.020]*	[.034]**	[.029]	[.029]**	[.026]	[.026]**
26–35	.152	.058	.050	.057	.042	.079	.047	.011	.062	.044	.090	.039	.097
	[.018]**	[.021]**	[.015]**	[.012]**	[.021]*	[.031]*	[.017]**	[.011]	[.017]**	[.023]$^{+}$	[.015]**	[.013]**	[.016]**

TABLE B1
(*Continued*)

Age Cohort	Chile	Czech Republic	Denmark	Finland	Germany	Hungary	Italy	Netherlands	Norway	Poland	Sweden	Switzerland	United States
36–45	.109	.071	.054	.037	.043	.069	.032	.017	.049	.032	.063	.050	.109
	[.024]**	[.018]**	[.014]**	[.012]**	[.029]	[.023]**	[.014]*	[.010]+	[.017]**	[.021]	[.019]**	[.014]**	[.016]**
46–55	.060	.079	.081	.042	.037	.044	.025	.015	.042	.012	.041	.006	.061
	[.024]*	[.020]**	[.015]**	[.014]**	[.026]	[.017]*	[.019]	[.013]	[.017]*	[.022]	[.017]*	[.018]	[.014]**
56–65	.032	.019	.090	.049	.018	.077	.036	.011	.106	.037	.078	.030	.068
	[.043]	[.023]	[.020]**	[.021]*	[.037]	[.028]**	[.023]	[.018]	[.025]**	[.022]+	[.021]**	[.026]	[.016]**
p-value (tertiary)	0	0	.05	.31	.05	.21	.14	.78	.01	.86	.06	.40	.04
Selectivity	.038	.498	.412	.629	.674	.301	−.194	.955	.364	.424	.431	.347	.309
	[.077]	[.092]**	[.064]**	[.070]**	[.107]**	[.106]**	[.105]+	[.068]**	[.098]**	[.072]**	[.080]**	[.065]**	[.098]**
IMR	−.626	−1.910	−1.947	−6.287	−4.888	2.482	.618	−8.715	−5.855	−1.893	−4.985	−5.926	−7.762
	[.833]	[1.434]	[2.334]	[2.442]*	[3.548]	[1.653]	[2.497]	[1.444]**	[2.644]*	[1.337]	[3.645]	[2.502]*	[1.915]**

Note.—Robust standard errors are in brackets. The sample includes individuals between 16 and 65 years of age, native born or immigrants completing at most primary education before immigration. The dependent variable is the normalized average literacy skill test score. Control variables are age, age squared, country-specific ability, country-specific IMR, country-specific indicators for mother's education, and a country-specific indicator for female. School's contributions to literacy skills for different cohorts are the coefficient estimates on the interactive terms between school education (measured by years of education at primary and secondary schools) and indicators for the respective age cohorts. College's contributions to literacy skills for different cohorts are the coefficient estimates on the interactive terms between college education (measured by years of education at tertiary schools) and indicators for the respective age cohorts. *p*-value (< tertiary) and *p*-value (tertiary) are for the *F*-tests that years of school less than tertiary and tertiary contribute the same over the five cohorts to the literacy score.

+ Significant at 10 percent.

* Significant at 5 percent.

** Significant at 1 percent.

TABLE B2
Estimation of Literacy Models for the United States with IQ Scores

	United States Only (1)	United States Only with IQ (2)	U.S. Results from Table 4 (3)
Age	.005	−.000097	
	[.015]	[.064]	
Age^2	0	.00015	
	[.000]	[.000]	
IQ		−.021	
		[.230]	
Female	.061	.061	.062
	[.031]$^+$	[.031]$^+$	[.031]*
Selectivity	.123	.123	.112
	[.094]	[.094]	[.093]
IMR	−12.383	−12.383	−10.463
	[1.883]**	[1.885]**	[1.811]**
School attainment by age cohort:			
16–25	.137	.137	.13
	[.012]**	[.012]**	[.011]**
26–35	.139	.139	.134
	[.010]**	[.010]**	[.010]**
36–45	.138	.138	.138
	[.010]**	[.010]**	[.010]**
46–55	.135	.136	.142
	[.009]**	[.010]**	[.009]**
56–65	.128	.128	.143
	[.011]**	[.011]**	[.010]**
p-value[a]	.76	.76	.49
Observations	2,392	2,392	
Adjusted R^2	.53	.53	

Note.—For col. 1, the *p*-value for the *F*-test of age = age^2 = 0 is 0. For col. 2, the *p*-value for the *F*-test of age = age^2 = IQ = 0 is .0001.
[a] The *p*-value is for the *F*-test that education's contributions to literacy score are the same over the five cohorts.
$^+$ Significant at 10 percent.
* Significant at 5 percent.
** Significant at 1 percent.

References

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Q.J.E.* 116, no. 1 (February): 313–50.

Ashenfelter, Orley, and Alan B. Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *A.E.R.* 84, no. 5 (December): 1157–73.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *J. Econ. Literature* 39, no. 4 (December): 1137–76.

Cameron, Stephen V., and James J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *J. Labor Econ.* 11, no. 1, pt. 1 (January): 1–47.

Card, David. 1999. "Causal Effect of Education on Earnings." In *Handbook of*

Labor Economics, edited by Orley Ashenfelter and David Card. Amsterdam: North-Holland, 1801–63.

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *J.P.E.* 100, no. 1 (February): 1–40.

Coleman, James S., et al. 1966. *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

Congressional Budget Office. 1987. *Educational Achievement: Explanations and Implications of Recent Trends*. Washington, DC: Congressional Budget Office.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: Elsevier, 697–812.

Dickens, William T., and James R. Flynn. 2001. "Heritability Estimates versus Large Environmental Effects: The IQ Paradox Resolved." *Psychological Rev.* 108 (2): 346–69.

Flynn, James R. 1984. "The Mean IQ of Americans: Massive Gains, 1932 to 1978." *Psychological Bull.* 95 (1): 29–51.

———. 1987. "Massive IQ Gains in 14 Countries: What IQ Tests Really Measure." *Psychological Bull.* 101 (2): 171–91.

———. 2007. *What Is Intelligence? Beyond the Flynn Effect*. Cambridge: Cambridge Univ. Press.

Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *J. Econ. Literature* 40, no. 2 (June): 436–82.

Hanushek, Eric A. 2002. "Publicly Provided Education." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein. Amsterdam: Elsevier, 2045–2141.

———. 2003. "The Failure of Input-Based Schooling Policies." *Econ. J.* 113, no. 485 (February): F64–F98.

Hanushek, Eric A., Victor Lavy, and Kohtaro Hitomi. 2008. "Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries." *J. Human Capital* 1, no. 2 (Spring): 69–105.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and the Estimated Effects of School Resources." *Rev. Econ. and Statis.* 78, no. 4 (November): 611–27.

Hanushek, Eric A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *J. Econ. Literature* 46, no. 3 (September): 607–68.

Harmon, Colm, Hessel Oosterbeek, and Ian Walker. 2003. "The Returns to Education: Microeconomics." *J. Econ. Surveys* 17 (2): 115–55.

Heckman, James J., Anne Layne-Farrar, and Petra Todd. 1996. "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings." *Rev. Econ. and Statis.* 78, no. 4 (November): 562–610.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North-Holland, 307–458.

———. 2008. "Earnings Functions and Rates of Return." *J. Human Capital* 2, no. 1 (Spring): 1–31.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *J. Labor Econ.* 24, no. 3 (July): 411–82.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Q.J.E.* 107, no. 1 (February): 35–78.

Lazear, Edward P. 2003. "Teacher Incentives." *Swedish Econ. Policy Rev.* 10 (3): 179–214.

Lowe, Roy, ed. 1992. *Education and the Second World War: Studies in Schooling and Social Change.* London: Falmer.

Lynn, Richard, and Tatu Vanhanen. 2002. *IQ and the Wealth of Nations.* Westport, CT: Praeger.

Maddison, Angus. 1995. *Monitoring the World Economy, 1820–1992.* Paris: OECD.

Mincer, Jacob. 1970. "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach." *J. Econ. Literature* 8, no. 1 (March): 1–26.

———. 1974. *Schooling, Experience, and Earnings.* New York: NBER.

Mulligan, Casey B. 1999. "Galton versus the Human Capital Approach to Inheritance." *J.P.E.* 107, no. 6, pt. 2 (December): S184–S224.

Murnane, Richard J., John B. Willett, Yves Duhaldeborde, and John H. Tyler. 2000. "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?" *J. Policy Analysis and Management* 19, no. 4 (Fall): 547–68.

Murphy, Kevin M., and Robert H. Topel. 1985. "Estimation and Inference in Two-Step Econometric Models." *J. Bus. and Econ. Statis.* 3, no. 4 (October): 370–79.

Murphy, Kevin M., and Finis Welch. 1992. "The Structure of Wages." *Q.J.E.* 107, no. 1 (February): 285–326.

Murray, T. Scott, Irwin S. Kirsch, and Lynn B. Jenkins, eds. 1997. *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey.* Washington, DC: Nat. Center Educ. Statis.

National Center for Education Statistics. 2005. *NAEP 2004: Trends in Academic Progress; Three Decades of Student Performance in Reading and Mathematics.* Washington, DC: U.S. Dept. Educ.

Neisser, Ulric, ed. 1998. *The Rising Curve: Long-Term Gains in IQ and Related Measures.* Washington, DC: American Psychological Assoc.

OECD (Organization for Economic Cooperation and Development). 1995. *OECD Education Statistics, 1985–92.* Paris: OECD.

———. 2005. *Education at a Glance: OECD Indicators 2005.* Paris: OECD.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22, no. 9 (September): 1325–44.

Psacharopoulos, George, and Harry A. Patrinos. 2004. "Returns to Investment in Education: A Further Update." *Educ. Econ.* 12, no. 2 (August): 111–34.

Smith, Jacqui, and Michael Marsiske. 1997. "Definitions and Taxonomies of Foundation Skills and Adult Competencies." In *Adult Basic Skills: Innovations in Measurement and Policy Analysis*, edited by Albert Tuijnman, Irwin S. Kirsch, and Daniel A. Wagner. Cresskill, NJ: Hampton.

Tamura, Robert F. 2006. "Human Capital and Economic Development." *J. Development Econ.* 79, no. 1 (February): 26–72.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2000. "Estimating the Labor Market Signalling Value of the GED." *Q.J.E.* 115, no. 2 (May): 431–68.

U.S. Office of Education. 1945. *Education under Enemy Occupation in Belgium, China, Czechoslovakia, Greece, Luxemburg, Netherlands, Norway, Poland.* Washington, DC: Fed. Security Agency, U.S. Office Educ.

Wolff, Richard J. 1992. "Italian Education during World War II: Remnants of Failed Fascist Education, Seeds of the New Schools." In *Education and the Second World War: Studies in Schooling and Social Change*, edited by Roy Lowe. London: Falmer.