The Causes and Consequences of Cross-Country Differences in Schooling Attainment*

Todd Schoellman †

Job Market Paper

December, 2006

Abstract

This paper uses labor market evidence to quantify the importance of quality-adjusted schooling differences in explaining cross-country income differences. I set up a model of equilibrium labor markets that is consistent with key facts about schooling attainment, education quality, and the returns to schooling within and across countries. The model suggests that the returns to schooling of immigrants to the U.S. are a measure of the education quality of their source country. Measured this way, quality differences across countries are large, and the calibrated model shows that schooling accounts for a factor of 5.9 of the income difference between the richest and poorest countries, and a factor of 2.67 between the top 20% of countries and the bottom 20%. These numbers are higher than was previously thought based on labor market evidence.

*Thanks to Robert Hall, Nir Jaimovich, Pete Klenow, Doireann Fitzgerald, John Taylor, Luigi Pisaferri, John Pencavel, Soohyung Lee, and Romans Pancs for many helpful comments. A particular thanks to Michèle Tertilt and Manuel Amador for their gracious guidance throughout this project. Any errors belong to me.

†Address: Department of Economics, Stanford University, Stanford, CA 94305. E-mail: todd.schoellman@stanford.edu.

1 Introduction

This paper uses labor market evidence to quantify the importance of quality-adjusted schooling differences in explaining cross-country income differences. It presents a model of worker schooling decisions under education quality differences that is consistent with three key facts about the way education quality interacts with other components of the labor market:

1. Countries with higher quality of schooling (as measured by international test scores) also have higher quantity of schooling (average years of schooling completed in the population).[1]

2. The log-wage returns to a year of schooling are weakly negatively correlated with the quality or quantity of schooling in a cross-section of countries.

3. The log-wage returns to a year of schooling for immigrants to the United States are strongly positively correlated with the education quality of their source country.

Traditional macroeconomic theories of cross-country schooling differences have trouble reconciling the second and third facts. In a standard setup, the log-wage returns to a year of schooling are independent of the country where the returns are observed.[2] If it were, then the returns to a year of Mexican education should be the same in Mexico or in the United States. In fact, the measured returns are 9.5 times higher in Mexico than in the United States. Conversely, the returns to a year of Swedish education are 2.9 times higher in the United States than in Sweden.

I set up a model of school quality that is consistent with facts 1, 2, and 3. The model has three key elements. First, workers make an endogenous decision about how to allocate their lifetime between schooling and labor. Going to school yields more human capital and higher future wages at the cost of foregone wages today. Workers go to school until the marginal log-wage returns to schooling equal their internal discount rate, as in Becker

[1]Similarly, Case and Deaton (1999) show that higher school quality has a positive effect on enrollment rates using micro data from South Africa, and Hanushek, Lavy, and Hitomi (2006) show the same for students in Egypt.

[2]Independence is a consequence of two common assumptions: that there is a single aggregate production function, and that labor services are provided in efficiency units. See, for instance, Mankiw, Romer, and Weil (1992), Hall and Jones (1999), Klenow and Rodriguez-Clare (1997), or Bils and Klenow (2000). The main analysis of Hendricks (2002) uses the efficiency units setup, but he also checks his analysis with a model that uses imperfect substitution across two skill types.

(1964). Note that as long as all workers have a similar internal discount rate, the marginal log-wage returns to a year of schooling will be constant across countries and unrelated to the quality or quantity of education in a country, so the model is consistent with fact 2.

The second key element is that the marginal returns to schooling are assumed to be increasing in the quality of education but decreasing in the level of schooling. Workers go to school until the marginal returns to schooling equal their internal discount rates, and workers across countries have similar discount rates. Hence, workers in high education quality countries go to school longer, and the model is consistent with fact 1; education quality and quantity are correlated across countries.

To resolve the discrepancy between facts 2 and 3, it is useful to decompose the returns to schooling into two sources. The returns to a year of schooling $R(S)$ are equal to the returns to a unit of human capital $R(H)$ times the human capital accumulated per year of schooling, which I define as the country's education quality Q:

$$R(S) = R(H)Q$$

Fact 2 says that $R(S)$ varies little across countries, but education quality differences imply that Q should vary a lot across countries. For these statements to be consistent, it must be the case that the return to a unit of human capital $R(H)$ varies inversely with Q, which is the third key element of the model. In high education quality countries, workers go to school longer and learn more per year of schooling, but the market returns to human capital adjust down. Looking at the returns to schooling across countries confounds the supply of human capital (Q) and the market price of human capital $R(H)$.

The model also makes predictions about the earnings of immigrants to the United States. They face a return to human capital that is consistent with the aggregate supply of human capital Q^{US} in the United States:

$$R^{US}(H) = \frac{R^{US}(S)}{Q^{US}}$$

The returns to a year of immigrant schooling in the U.S. is then:

$$R^{IMM}(S) = R^{US}(H)Q^{IMM} = R^{US}(S)\frac{Q^{IMM}}{Q^{US}}$$

The U.S. labor market returns to a year of foreign schooling are directly proportional to the quality of the foreign schooling. The model is consistent with fact 3: immigrants from high education quality countries earn higher returns on their schooling in the U.S. Further, the model allows me to use the log-wage returns to schooling of immigrants to the United States as a measure of the education quality of their source country.[3] Looking at immigrants to the United States fixes the labor market demand conditions (the returns to human capital) and makes it possible to draw inferences about the labor market supply of immigrants (the human capital per year of schooling).

The model is capable of qualitatively matching the three key facts outlined in the opening paragraph. I calibrate the model to match these facts quantitatively. I use the marginal log-wage returns to schooling of immigrants to the U.S. to measure the education quality of 130 foreign countries. I fit the model to data on schooling variation and the returns to schooling within and across countries. The calibrated model is capable of fitting these data quite closely; in particular, the immigrant education quality measure leads to reasonable variation in schooling levels.

It is then possible to back out of the calibrated model the implied importance of cross-country schooling differences for explaining income differences. Schooling differences account for a factor of 5.9 of the observed income difference between the richest and poorest countries, as opposed to the usual factor of 2-3.[4] Schooling differences account for a factor of 2.67 of the income differences between the top 20% and the bottom 20% of countries.

I check the robustness of the model to including worker heterogeneity through local school quality or ability differences. Heterogeneity opens the door to discussing two important points. First, the observed Mincerian returns may not be the true private returns to schooling, but may also incorporate some sorting of high-ability workers into high levels of schooling.[5] Second, immigrants may selected and may be different from non-migrants.[6]

[3]Here, this paper draws on previous research that estimates education quality from labor market returns to schooling, particularly Card and Krueger (1992) and Bratsberg and Terrell (2002). Bratsberg and Terrell use immigrant returns to measure source country education quality in a paper that investigates the determinants of education quality. Using returns in this way is controversial since Heckman, Layne-Farrar, and Todd (1996); I attempt to incorporate and address some of their concerns.

[4]For instance, Hall and Jones (1999) report a factor of 2.5.

[5]There is a long labor literature on this point. See Griliches (1977) for an early review and discussion. Card (1999) reviews the recent literature with a strong focus on the different instrumental variables strategies that have been used. Heckman, Lochner, and Todd (2006) includes a critique of the instrumental variables approach in this context.

[6]Borjas (1987) and Borjas (1999) are useful references for the immigrant selection problem. Bratsberg and Terrell (2002) control their variables for the classic problem of selection in unobserved ability. Here, I am concerned with a different problem: it may be that the education quality of immigrants is estimated

I suggest a structural way to measure the implied immigrant selection, and find evidence that immigrants from Africa in particular tend to be most selected, while immigrants from Latin America tend to be least selected. Accounting for possibility that Mincerian returns may overstate private returns to schooling lowers the predicted importance of schooling. The model accounts for a factor of 2.91 of the income difference between the richest and the poorest countries, and a factor of 1.89 between the top 20% and the bottom 20%.

Other recent research has advocated the importance of accounting for school quality. Bils and Klenow (2000) allow the value of schooling to be affected by the human capital of previous generations. Their mechanism is essentially one of school quality which is generated by teacher quality. Manuelli and Seshadri (2005) and Erosa, Koreshkova, and Restuccia (2006) allow workers to create human capital by investing time and market resources, as in Ben-Porath (1967). Market resources can be interpreted as buying school or human capital quality. These three papers have the advantage of specifying exactly what mechanisms may lead to school quality and how the mechanisms affect the measured role of schooling. However, empirical research on the determinants of school quality remains skeptical about the quantitative importance of schooling expenditures or the role of teacher education (Hanushek 1995, Hanushek 2002). Hence, it is useful to explore a framework which measures general school quality differences without specifying the channels through which they arise.

The paper proceeds as follows. Section 2 presents the data and gives the three key relationships in more detail. Section 3 provides a model with ex-ante identical workers that is consistent with these facts and provides interpretation for them. Section 4 calibrates the model and measures the income predictions. Section 5 shows how a model with heterogeneous workers leads to different interpretations of the data. Section 6 measures the implied immigrant selection and re-calibrates the model. Section 7 concludes.

2 Cross-Country Data on Schooling

2.1 Four Sources of Data

The three key facts from the introduction are built on data about education quantity, education quality, and the returns to a country's education at home and abroad. Here I introduce the underlying data sources and the facts in more detail. The data on schooling

correctly, but that immigrants have education quality which is atypical for non-migrants.

attainment is average years schooling in the over-25 population of different countries in 1999, taken from the Barro-Lee data set (Barro and Lee 1996, Barro and Lee 2001).

I use data on international test scores from three different sources as a measure of education quality across countries. Hanushek and Kimko (2000) construct a test score index for a broad cross-section of countries by aggregating a series of different testing programs running from 1966-1991. The OECD Programme for International Student Assessment (PISA) administered exams to 15 year-olds enrolled in school in a large number of countries in 2000 and 2003. The U.S. Department of Education Trends in International Mathematics and Science Study (TIMSS) administered a similar exam in 1995, 1999, and 2003 to 13 year-olds enrolled in school. I use test score data as a consistency check on the immigrant-based measures of education quality. Appendix A provides further details on the the data used and how the test score indices are constructed.

Returns to schooling are commonly measured using Mincerian returns gathered from a regression suggested in Mincer (1974):

$$\log(W^k) = b + MS^k + \beta X^k + \varepsilon^k \tag{1}$$

where W^k is the wage of individual k, S^k is years of schooling, X^k is a vector of controls, and ε^k is the error term. M is the Mincerian returns to schooling. Interpreting M as the private returns to schooling is controversial, since unobserved ability may bias up both wages and schooling attainment; see the papers cited in the introduction for more detail. For now, I present the data on M collected from OLS regressions of the form of equation (1) without taking a stand on what they represent. I consider models where M is the private returns and where it overstates the private returns.

Psacharopoulos and a coauthor (Psacharopoulos 1994, Psacharopoulos and Patrinos 2004) have collected observations on the Mincerian returns to schooling for a number of different countries.[7] Each observation is an estimate of M from equation (1) taken from an outside study. I use the update on this data set provided by Banerjee and Duflo (2005), who add some newer observations and remove some others based on studies judged to be of low quality by Bennel (1996).

The fourth source of data is the returns to schooling of immigrants to the United States.

[7] Trostel, Walker, and Woolley (2002) have also collected Mincerian returns for a small sample of countries. Their returns have the benefit that they are collected from more similar underlying data. The general patterns there are similar to the ones emphasized here.

This approach follows Card and Krueger (1992), who use the returns to schooling of cross-state migrants to identify the quality of education of the migrants' source state. They run a regression of the form:

$$\log(W_j^{i,k}) = b^i + b^j + M^i S_j^{i,k} + \beta X_j^{i,k} + \varepsilon_j^{i,k} \tag{2}$$

where $W_j^{i,k}$ is the wages of individual k who received schooling in state i and then migrated to and works in state j. The authors control for state of origin and state of residence fixed effects b^i and b^j, and interpret the slope of wages with respect to schooling M^i as a measure of state i's education quality.

Bratsberg and Terrell (2002) adopt this methodology to study the education quality of other countries by examining the log-wage returns to schooling of immigrants to the U.S. I repeat their exercise with a few changes using 2000 U.S. Census data. The regression equation is:

$$\log(W_{US}^{i,k}) = b^i + M_{US}^i S_{US}^{i,k} + \beta X_{US}^{i,k} + \varepsilon_{US}^{i,k} \tag{3}$$

where $W_{US}^{i,k}$ now represents the earnings of immigrant k from country i in the United States.

I implement this equation using the 5% sample of the 2000 Census Public Use Micro Survey, made available through the IPUMS system (Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander 2004). Immigrants are identified by country of birth. The Census lists separately each of 130 statistical entities with at least 10,000 immigrants counted in the United States. Some of these statistical entities are nonstandard: for instance, there are response categories for Czechoslovakia, the Czech Republic, and Slovakia, since immigrants came both before and after the split. I preserve every statistical entity which is separately identified, and refer to them as countries as a shorthand.[8]

The Census includes a measure of schooling attainment which I recode as years of schooling in the usual manner. The Census does not provide direct information on where the schooling was obtained. Instead, I use information on age, year of immigration, and schooling attainment to impute which immigrants likely completed their schooling before immigrating. It is important to exclude from the sample immigrants who may have received some or all of their education within the United States to have an unbiased estimate of

[8]There are two exceptions to this rule: I exclude the USSR, Russia, and North Korea because the data do not allow me to differentiate these groups to the extent necessary. I also merge the United Kingdom into a single observation.

source-country education quality. The sample used includes all workers who immigrated to the United States at least six years after their expected date of completing their reported education. For instance, a high school graduate (12 years of schooling) would need to be at least 24 (start school at 6, go to school 12 years, add 6) to be included in the sample. The extra six years are allowed to help minimize the noise coming from workers who repeat grades or delay school, for instance to fulfill mandatory armed forces obligations.[9]

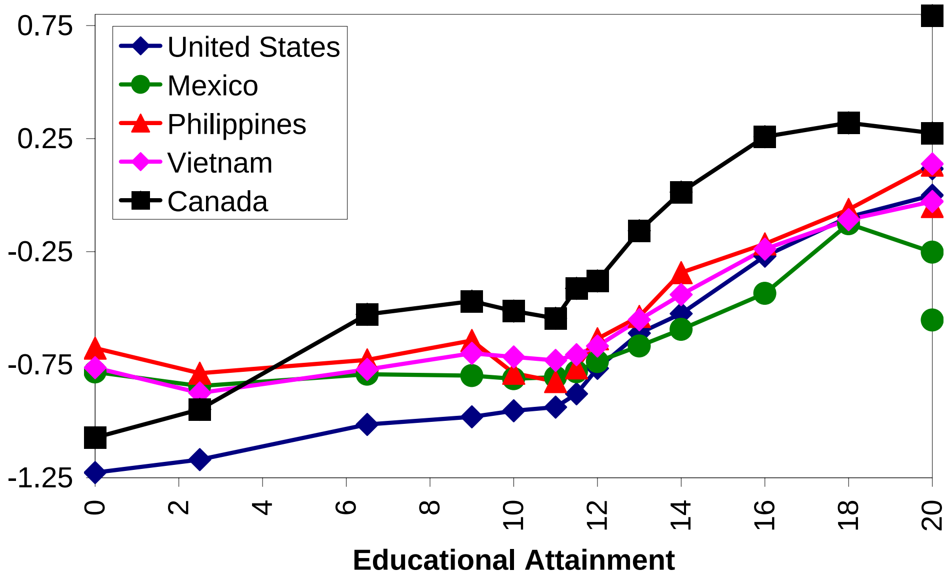
The sample includes respondents aged 19-64 who were employed but not self-employed in the previous year. The wage is calculated as weekly wage by dividing the previous year's wage income by weeks worked in the previous year. The vector of controls includes age and its square, gender, a dummy for residence in metropolitan area, self-assessed English language proficiency on a five-option scale, dummies for Census region of residence, a disability dummy, and a full set of year of immigration dummies.

The final sample includes 4.1 million Americans and 220,000 immigrants from 130 different source countries. Appendix C provides regression results, including number of observations and standard errors. The measured U.S. returns are 10.2%. Results from other countries vary widely. The highest returns are observed for immigrants from Tanzania, Sweden, and Belgium, at 14-15%. A few countries have negative returns, although none of these coefficients is statistically significant. Two useful benchmarks on the low end of the scale are Laos and Mexico, with 0.8% and 0.9% returns estimated with a large sample of immigrants.

The underlying wage and education patterns are also of some interest. I re-estimate (3), replacing the linear schooling effect $M^i_{US} S^{i,k}_{US}$ with an unrestricted vector of dummies $M^i_{US}(E) S^{i,k}_{US}(E)$ for fourteen different educational attainment levels E. Figure 1 plots $M^i_{US}(E)$ against the typical years of schooling $S(E)$ for the United States and four of the countries with the highest immigrant counts in the U.S. Immigrants with no education are always paid more than Americans. As the education level increases, the earnings of natives overtake those of immigrants from some countries (Mexico) but fail to catch up to those of others (Canada).

Heckman, Layne-Farrar, and Todd (1996) argue that the Card and Krueger (1992) cross-state education quality differences are driven entirely by different quality of college education. Their work suggests that the returns to the first twelve years of education are

[9]Failing to allow for interruptions leads to returns that are higher on average, which I interpret as evidence that they are biased by including some portion of source country returns and some portion of the high U.S. returns. Allowing for larger interruptions reduces the sample size without significantly changing the average returns.



Note: Educational attainment levels are taken from the 2000 U.S. Census data. They are: less than first grade, first-fourth, fifth-eight, ninth, tenth, eleventh, twelfth, twelfth with diploma or GED, some college, Associate, Bachelors, Doctorate, and Professional.

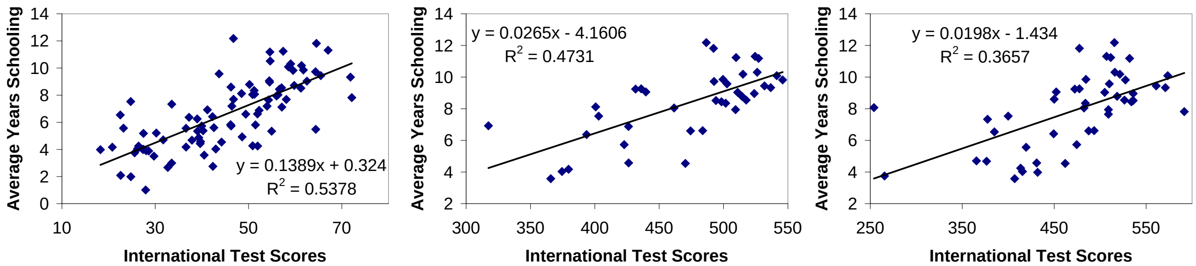
Figure 1: Wage by Educational Attainment, Relative to U.S.

nearly identical across states. The same fact does not apply across countries, which can be seen by comparing the slopes of the wage profiles of different countries. The Mexican profile is generally flatter than the Canadian profile at every educational level, indicating lower returns to every education level, not just college graduation. These findings are consistent with the results from test score measures of primary and secondary school students, which also suggest large quality differences throughout the education process.

2.2 Key Data Relationships

I can now present the three key data relationships in more detail. Fact 1 is that average schooling attainment and quality of schooling are positively correlated across countries. Figure 2 plots average years schooling and test scores for each of the three test score measures; the relationship is positive and significant for each measure.

Fact 2 is that the Mincerian returns to schooling in a country are uncorrelated with school quantity and school quality. The relationship between a country's average years of schooling and its Mincerian returns has received a great deal of attention in the literature, so I present it separately. The data collected in Psacharopoulos (1994) suggested that low-



(a) Hanushek-Kimko (b) PISA (c) TIMSS

Figure 2: Relationship Between School Quality and School Quantity

schooling countries had higher observed Mincerian returns; this relationship was generally interpreted as evidence of diminishing returns to schooling, and was incorporated as such by Bils and Klenow (2000). Figure 3 plots the updated Banerjee and Duflo (2005) Mincerian returns data against average years of schooling for the cross-section of countries with data available for both series. As noted by Banerjee and Duflo, the relationship between average years of schooling and the returns to schooling is much weaker once low-quality studies are removed from the data set. The point estimate is statistically significant but represents a small fraction of the total variation.

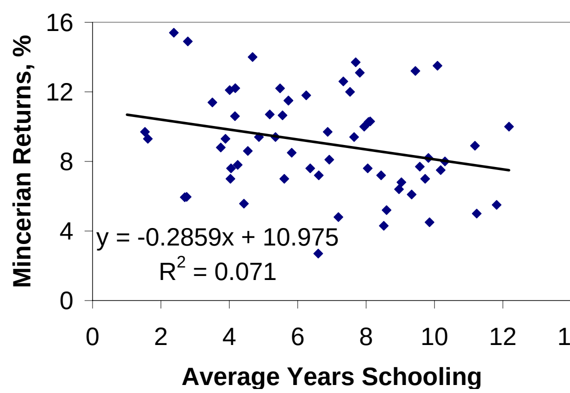
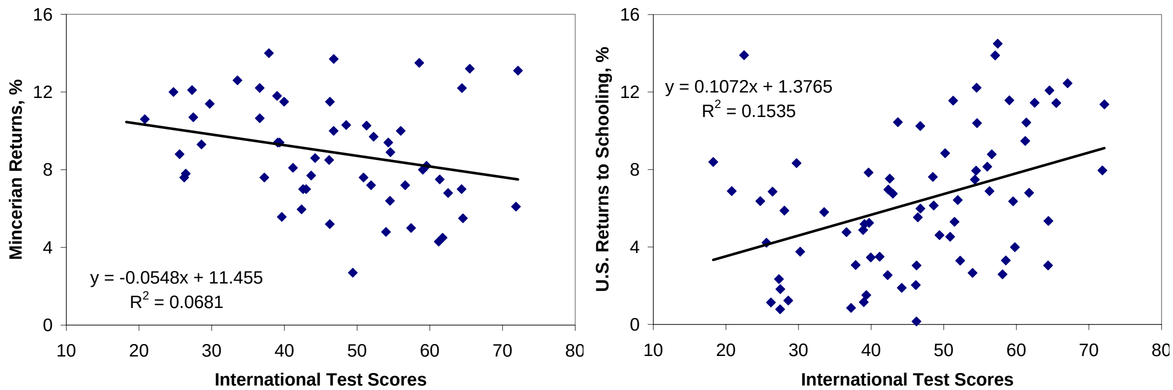


Figure 3: The Schooling-Returns Relationship

Finally, I emphasize the relationship between quality and returns. Here, returns to country j schooling can be observed in two locations: among non-migrants in country j, and among immigrants to the United States. Facts 2 and 3 say that the returns observed in country j are uncorrelated with j's education quality, but the returns observed in the

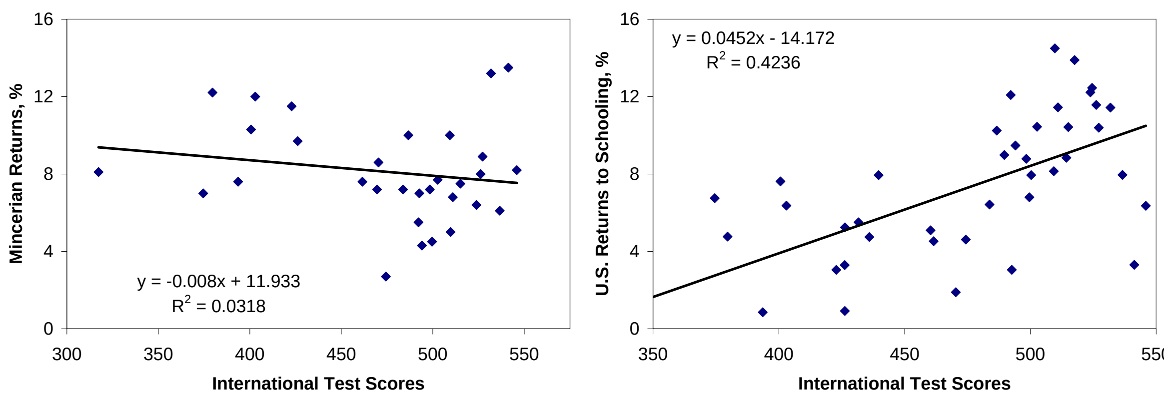
United States are positively correlated with j's education quality. Figures 4 - 6 plot both types of returns for the three test score measures of quality. Returns observed in the same country are actually slightly negatively correlated with education quality.



(a) Own-Country Returns

(b) U.S. Immigrant Returns

Figure 4: Hanushek-Kimko Quality and Returns to Schooling

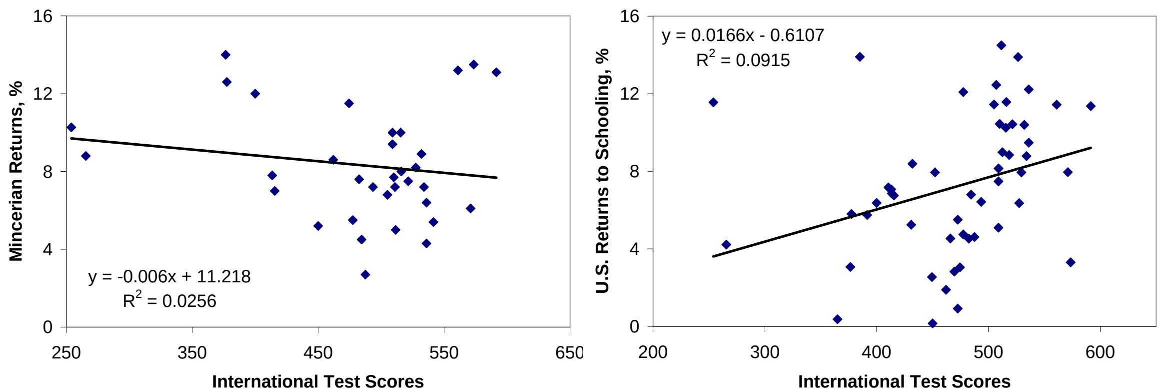


(a) Own-Country Returns

(b) U.S. Immigrant Returns

Figure 5: PISA Quality and Returns to Schooling

These are the three key facts of the paper. I now turn my attention to writing down a model which qualitatively replicates and interprets these facts.



(a) Own-Country Returns

(b) U.S. Immigrant Returns

Figure 6: TIMSS Quality and Returns to Schooling

3 A Model with Identical Workers

3.1 Introduction to Model

The heart of the model is the labor market for workers of different education levels across countries. Equilibrium in the labor market is determined by the interaction between employers and workers. Employers are drawn from a set of industries with different demands for skills. Some industries produce landscaping services and have technologies which are unintensive in skills, while other industries produce new pharmaceuticals and have technologies which are intensive in skills. Workers take the education quality and the set of employers as given and choose how long to go to school and how long to work. The labor market equilibrium includes the level of schooling for workers, the returns to schooling of workers, and the returns to schooling of immigrants.

3.2 Population Dynamics

The world consists of J closed economies, with individual economies indexed by subscript j. Each economy has a continuum of measure 1 of identical workers, with a representative worker denoted by d. Workers lose their human capital every T_j years, with the date of human capital loss staggered across workers so that exactly T_j^{-1} workers forget their human capital each period. The setup is identical to one in which each d is a dynasty of altruistically linked workers who live T_j years.[10]

3.3 Dynasties

Each dynasty in this model has the same power felicity function with intertemporal elasticity of substitution σ. Preferences over sequences of consumption $c_j(d,t)$ are given by:

$$U = \int_0^\infty e^{-\rho t} \frac{c_j(d,t)^{1-1/\sigma} - 1}{1 - 1/\sigma} dt$$

where ρ is the usual rate of time discounting.

At time t, the dynasty has two sources of income: labor income $\omega_j(d,t)$, to be defined in the next section, and the returns on asset holdings, $R_j(t)a_j(d,t)$. The dynasty spends this income on consumption $c_j(d,t)$ and on changes in its net asset holdings $\dot{a}_j(d,t)$. Then its period budget constraint is given by:

$$c_j(d,t) + \dot{a}_j(d,t) = \omega_j(d,t) + R_j(t)a_j(d,t) \tag{4}$$

The dynasty's preferences can be summarized by the standard Euler equation:

$$\frac{\dot{c}_j(d,t)}{c_j(d,t)} = \sigma(R_j(t) - \rho) \tag{5}$$

3.4 Labor-Schooling Decision

Time spent in schooling creates human capital and raises future wages. A worker with $S_j(d,t)$ years of schooling has human capital given by the human capital production function:

$$H_j(d,t) = \exp\left[\frac{(S_j(d,t)Q_j)^\eta}{\eta}\right] \tag{6}$$

This function is based on the one used in Bils and Klenow (2000). I assume $0 < \eta < 1$ throughout the paper, so that there are diminishing log-marginal returns to schooling. Country j's education quality level is given by Q_j, which determines the human capital

[10] Altruistically linked in the standard sense of Barro (1974). The only demographic factor accounted for here is average life expectancy, but Manuelli and Seshadri (2005) also find large effects when they account for differences in the age distribution of the population. I also ignore issues related to stochastic mortality; see Kalemli-Ozcan, Ryder, and Weil (2000), Soares (2005), or Cervellati and Sunde (2005) for a model where this uncertainty may be important.

acquired per year of schooling. Higher Q_j offsets some of the diminishing returns to schooling.

Workers are endowed with one unit of time each period. To make the problem simple, I abstract from any financial costs of schooling and from the possibility that workers may prefer time spent in schooling to time spent in the labor force. Then workers take the wage as a function of schooling $W_j(S,t)$ as given and allocate their time between schooling and labor to maximize lifetime income. As is standard, workers separate their lives into two periods: they go to school full-time from the beginning of their life until age S, after which they work full-time. The problem of a worker born at time τ is to choose $S_j(d,t)$ to maximize income:

$$\omega_j(d,t) = \max_{S_j(d,t)} \int_{\tau+S_j(d,t)}^{\tau+T_j} e^{-\int_\tau^t R_j(\theta)d\theta} W_j(S_j(d,t),t)dt$$

The worker's income maximization problem has first order condition

$$\int_{\tau+S_j(d,t)}^{\tau+T_j} e^{-\int_\tau^t R(\theta)d\theta} \frac{\partial W_j(S_j(d,t),t)}{\partial S_j(d,t)} dt = e^{-\int_\tau^{\tau+S_j(d,t)} R(\theta)d\theta} W(S_j(d,t), S_j(d,t)+\tau) \quad (7)$$

Workers go to school until the present discounted value of the future wages gained from an additional unit of schooling (left-hand side) equals the wage foregone by obtaining that unit of schooling (right-hand side).

3.5 Intermediate Goods Producers

Each country has a continuum of industries distributed uniformly on $[\underline{\gamma},\bar{\gamma}]$. Each industry produces a unique intermediate good. Industries vary in the skill-intensity of their production process. Output in industry γ is given by the production function:

$$Y_j(\gamma,t) = K_j(\gamma,t)^\alpha \left(A_j(t)H_j(\gamma,t)^\gamma L_j(\gamma,t)\right)^{1-\alpha} \quad (8)$$

where $A_j(t)$ is the labor-augmenting efficiency level general to the entire country, $K_j(\gamma,t)$ and $L_j(\gamma,t)$ are the capital and labor choices specific to the industry, and $H_j(\gamma,t)$ is the human capital per worker. Efficiency grows exogenously at rate g. Given the human capital accumulation function in equation (6), higher γ industries are more skill-intensive. The standard analysis using human capital assumes $\gamma = 1$ (Bils and Klenow 2000, Hall and Jones 1999). I explore the variation that arises as skill intensity varies around the mean of

1.

There is a large set of potential entrants, so that no profits are earned and the equilibrium number of firms in each industry is indeterminate. Then industry γ takes the time t real price of its output $P_j(\gamma, t)$, the rental price of capital $R_j(t) + \delta$, and the schedule of wages $W_j(S, t)$ as given. It chooses the capital stock, labor supply, and level of schooling per worker to maximize profits each period. Substituting in for human capital, the industry's problem is:

$$\max_{K_j(\gamma,t),L_j(\gamma,t),S_j(\gamma,t)} P_j(\gamma,t)K_j(\gamma,t)^{\alpha}\left\{A_j(t)\exp\left[\frac{\gamma}{\eta}(S_j(\gamma,t)Q_j)^{\eta}\right]L_j(\gamma,t)\right\}^{1-\alpha}$$
$$-(R_j(t)+\delta)K_j(\gamma,t) - W_j(S_j(\gamma,t),t)L_j(\gamma,t)$$

The first-order conditions for an interior solution are:

$$\alpha P_j(\gamma,t)\frac{Y_j(\gamma,t)}{K_j(\gamma,t)} = R_j(t)+\delta \tag{9}$$
$$(1-\alpha)P_j(\gamma,t)\frac{Y_j(\gamma,t)}{L_j(\gamma,t)} = W_j(S_j(\gamma,t),t) \tag{10}$$
$$(1-\alpha)\gamma P_j(\gamma,t)S_j(\gamma,t)^{\eta-1}Q_j^{\eta}\frac{Y_j(\gamma,t)}{L_j(\gamma,t)} = \frac{\partial W_j(S_j(\gamma,t),t)}{\partial S_j(\gamma,t)} \tag{11}$$

Combining (10) and (11) yields the equation that relates the optimal schooling level to the log-wage returns to schooling:

$$\frac{\partial \log(W(S_j(\gamma,t),t))}{\partial S_j(\gamma,t)} = \gamma S_j(\gamma,t)^{\eta-1}Q_j^{\eta} \tag{12}$$

3.6 Balanced Growth Path

A final goods producer uses a CES production function with elasticity of substitution ψ to aggregate the intermediate goods into a final good suitable for consumption or investment. Since the focus here is on the interaction between intermediate producers and workers, the firm's problem is described in Appendix B. I also define an equilibrium and the balanced growth path there. For the rest of the paper I confine my attention to the balanced growth path. Most of the equations simplify.

The real interest rate is constant over time and across countries at $R = \frac{g}{\sigma} + \rho$. The optimal schooling decision of workers is:

$$\frac{R-g}{1-\exp[-(R-g)(T_j-S_j(d))]} = \frac{\partial \log(W_j(S_j(d)))}{\partial S_j(d)} = M_j$$

where M_j is the Mincerian returns to schooling in country j, as discussed in Section 2. For ease of exposition, I adopt the additional assumptions that $\sigma = 1$, i.e. that the felicity function is log, and that the equilibrium $T_j - S_j$ is large. The second assumption is standard in the labor literature, but somewhat unusual here where schooling is endogenous. I use it only to present simplified results familiar from that literature; I drop both assumptions when I calibrate the model. Under these assumptions, the Mincerian returns are given by:

$$M_j = R - g = \rho \tag{13}$$

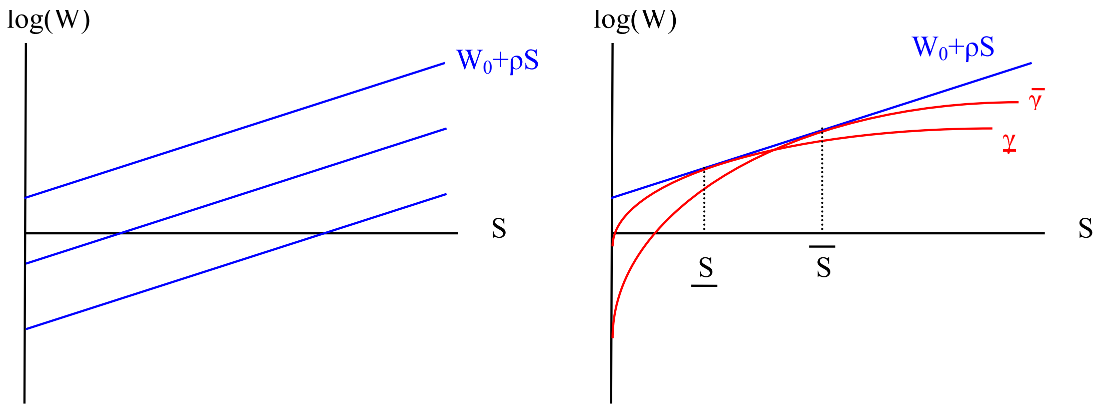
Since ρ is a constant parameter across countries, the model is consistent with fact 2: Mincerian returns are uncorrelated with quantity or quality of schooling.

Equation (13) is from Becker (1964): workers go to school until the marginal log-wage returns to schooling equal their internal discount rate. As in Mincer (1958), this condition does not define the optimal schooling decision of a worker. Rather, since all workers are ex-ante identical, it defines the indifference curve of the representative worker: a worker is willing to get any number of years of schooling, as long as he receives an appropriately higher wage upon graduation. Figure 7a draws their indifference curves, which are linear when plotted in log-wage-schooling space, with higher indifference curves representing higher welfare.

The behavior of the intermediate industries also simplifies along the balanced growth path. Rearranging their first-order conditions yields the zero-profit log-wage schedule offered in industry γ:

$$\log(W_j(S,t;\gamma)) = A(0) + gt + \log\left[\frac{(1-\alpha)\alpha^{\alpha/(1-\alpha)}P_j(\gamma)^{1/(1-\alpha)}}{(R+\delta)^{\alpha/(1-\alpha)}}\right] + \frac{\gamma}{\eta}(SQ_j)^{\eta} \tag{14}$$

The log-wage has an intercept whose value depends on aggregate efficiency and the industry output price $P_j(\gamma)$, and whose slope depends on the industry skill-intensity γ. In equilibrium, all varieties are produced, so workers must be willing to work in any industry. Industry prices $P_j(\gamma)$ adjust to satisfy worker indifference. The equilibrium is shown in Figure 7b for two industries, the least ($\underline{\gamma}$) and the most ($\bar{\gamma}$) skill-intensive. Less skill-intensive



(a) Workers' Indifference Curves

(b) Equilibrium Labor Market

Figure 7: Labor Market

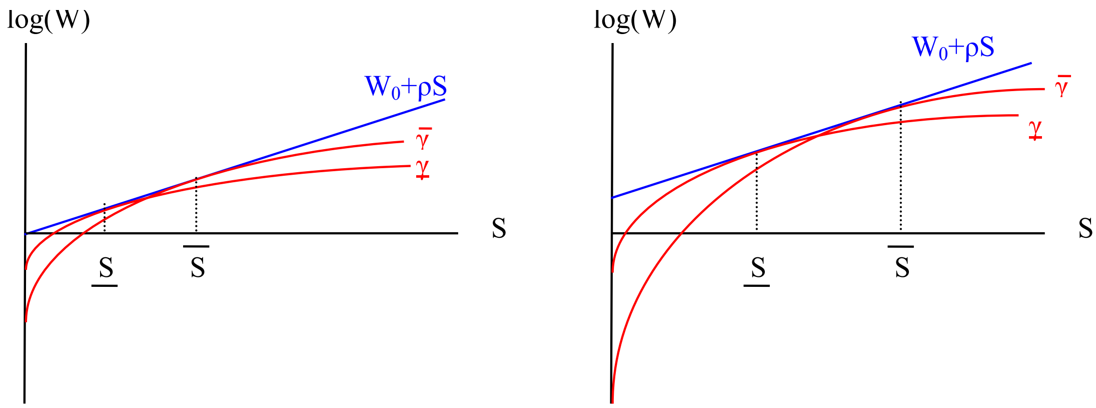
industries have higher prices and post wage schedules with a higher fixed component but a lower return to schooling. The equilibrium level of schooling is given by the point of tangency between workers' indifference curve and the industries' zero-profit wage schedule:

$$S_j(\gamma) = \left[\frac{\gamma Q_j^\eta}{\rho}\right]^{1/(1-\eta)} \tag{15}$$

Given $0 < \eta < 1$, higher education quality leads to higher average school attainment, consistent with fact 1.

Figure 8 shows how labor markets differ for countries with different education quality. Two effects are straightforward: workers in the high quality country accumulate more schooling and are on a higher indifference curve than the workers in the low quality country. A third effect is more subtle. In the high education quality country the intercepts of the wage contracts are further apart. From (14) the larger difference between intercepts is an indication of larger differences in the prices of the goods.

Adjustment in the prices of the intermediate goods is the key mechanism that keeps Mincerian returns constant across economies with very different education quality. In low education quality countries, the relative price of high γ goods to low γ goods is close to 1. As education quality increases, workers with more schooling should tend to earn a higher wage premium, but the quality effect is offset by a decrease in the relative prices of the goods they produce. In equilibrium, prices adjust so that the Mincerian returns are always equal to ρ.



(a) Low-Quality Country (b) High-Quality Country

Figure 8: Labor Markets in Two Different Countries

Other work in this literature has stressed the importance of being consistent with observed cross-country Mincerian returns. However, papers consistent with these returns have produced a wide variety of estimates for the importance of schooling in accounting for income differences. At the extremes, Hall and Jones (1999) find that schooling accounts for a factor of 2.5 income difference, while Manuelli and Seshadri (2005) find that schooling accounts for a factor of 20 income difference. Observed Mincerian returns are consistent with a variety of labor market theories. Here, optimizing workers ensure that returns are constant across countries regardless of the value of schooling. Mincerian returns do not discriminate sufficiently between alternative theories of the role of human capital in explaining cross-country income differences, so I use different measures derived from the labor market performance of immigrants.

Suppose that a worker immigrates from another country i with schooling attainment S_j^i, and that country i has education quality Q_i. As long as $Q_i \neq Q_j$, this worker will earn a different wage than a native worker with the same schooling attainment because they will have different human capital levels. The slope of his log-wage schedule is given by:

$$M_j^i = \frac{Q_i}{Q_j} M_j \qquad (16)$$

The returns to foreign schooling are directly proportional to the relative education quality. Immigrants from countries with education quality half of the domestic level accumulate half as much human capital per year of schooling and earn half the rate of return per year of schooling. The model is consistent with fact 3.

Equation (16) also says that the returns to schooling of immigrants are a direct measure of their source country education quality. In the next section, I use these measures of quality as an input to the model. I calibrate it to fit facts 1, 2, and 3 quantitatively, and then study the implied role of schooling in accounting for observed cross-country income differences.

4 Calibration of the Model

4.1 Fixing Model Parameters

Calibrating the model requires a set of J countries with the necessary inputs to the model: $(\{Q_j, T_j, A_j\}_{j=1}^{J})$. Values of Q_j are taken from returns to immigrant schooling; see Table 6. T_j, the potential working life span in country j, is taken to start at age 5 and continue until the average worker dies or retires. Hence, I set T_j equal to the country's age 5 life expectancy (worker dies before expected retirement) or 60 (retirement at 65), whichever is smaller.[11] A_j is set equal to one, which is innocuous in this model since efficiency does not affect schooling decisions. Then the income differences predicted by the model come solely from differences in education and life expectancy. I compare the performance of the calibrated model against the actual data values S_j and M_j, which are taken from the data sources listed in Section 2. Finally, I compare the model's income predictions to the data values for Y_j, measured as PPP income per capita from the WDI data (World Bank 2006).

Relatively few countries have all 6 data points. Rather than discard a large fraction of the sample, I aggregate over countries. I take the 117 countries that have data for both Y_j and Q_j and form them into five quintiles by income per capita. For each quintile, the values $(Q_j, T_j, S_j, M_j, Y_j)$ are the average values for the quintile, ignoring missing observations. Table 1 gives the resulting values for the five quintiles, as well as the United States for reference.

The model also requires ten parameters to be fully specified. Using immigrant returns identifies education quality relative to the U.S. level, so I need to calibrate one normalization $\bar{Q}$; then each country's education quality is $Q_j = M_{US}^{j}\bar{Q}$. The other 9 parameters are:

[11] Age 5 life expectancy is estimated using data on life expectancy at birth, infant mortality rate, and under 5 mortality rate, taken from the 2005 Human Development Report (United Nations Development Programme 2005). I assume that infants who die before age 1 live 0.25 years on average, and that those who die between ages 1 and 5 live 3 years on average.

Table 1: Representative Quintiles

	Country					
Observation	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	US
M_{US}^j	4.14%	4.02%	4.39%	5.73%	9.70%	10.2%
M_j	8.72%	10.0%	11.8%	8.76%	8.06%	10.0%
Life Exp.	61.4	72.5	72.8	74.4	79.5	77.9
PPP GDP p.c.	1,545	4,004	6,510	13,356	28,060	36,465
Schooling	2.80	5.09	5.89	7.58	9.46	12.2

[a] M_{US}^j and M_j are the log-wage returns to schooling of immigrants to the U.S. and non-migrants. Life expectancy is age 5 life expectancy calculated as explained in the text; the value used for T_j is different, as reported. Schooling is average years schooling from Barro-Lee.

$\eta, \underline{\gamma}, \bar{\gamma}, \alpha, \rho, g, \delta, \sigma$, and ψ. Values for some of these parameters are based on outside evidence or convenient benchmarks. $\alpha = 0.33$, $\sigma = 0.5$, and $\delta = .06$ are standard values from the real business cycle literature (Cooley and Prescott 1995). I set $g = 1.75\%$ to match the average long-term growth rate of real GDP/capita for a large sample of countries.[12] The distribution $[\underline{\gamma}, \bar{\gamma}]$ is centered on $\gamma = 1$ to make my work more comparable with the existing literature.

Then five parameters need to be calibrated in the model: ρ, η, $\bar{\gamma}$, $\bar{Q}$, and ψ. I calibrate the model using a set of moments related to each parameter's role in the model. ρ determines the Mincerian returns, so the first moment is the world average Mincerian returns of 9.1%. The model predicts little variation across countries, so trying to match each country's returns yields little additional benefit.

$[\underline{\gamma}, \bar{\gamma}]$ determines within-country schooling variation between the most and least skill-intensive industries. The 2000 U.S. Census organizes workers into 475 occupation codes. The average years of schooling by occupation for employed natives ranges from 10.93 years of schooling for dishwashers to 19.87 years for optometrists. I compare the model and the data on the basis of the ratio between these moments, so the second moment $\frac{S_{US}(\underline{\gamma})}{S_{US}(\bar{\gamma})} = \frac{19.87}{10.93} = 1.82$.

$\bar{Q}$ determines the average world schooling level and η determines the cross-country variation in schooling that arises from observed education quality differences. I use the

[12]Barro and Sala-i Martin (1999), p.3.

schooling levels of the five quintiles as the corresponding moments for these parameters.

The elasticity of substitution across varieties ψ is difficult to pin down directly because varieties are defined differently here than is standard in the trade literature, for instance. However, elasticity across varieties also implicitly determines the elasticity of substitution across workers with different education levels. Katz and Murphy (1992) estimate that the elasticity of substitution between high school and college educated workers in the United States is 1.4. I use this moment to calibrate ψ.

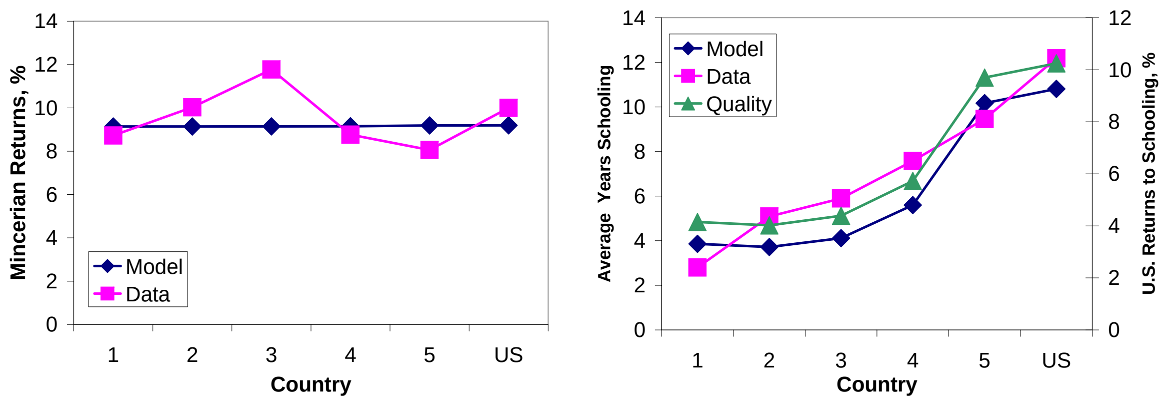
Table 2: Baseline Model Calibrated Parameters

Parameter	Role	Value
	Calibrated to Outside Evidence	
α	Capital Share	0.33
δ	Capital Depreciation Rate	0.06
σ	Intertemporal Elasticity of Substitution	0.5
g	GDP p.c. Growth Rate	1.75%
	Calibrated to Fit Data	
ρ	Time Discount Rate	0.073
η	$\varepsilon_{H,S}$	0.54
$\bar{Q}$	Quality Level	0.92
$\underline{\gamma}$	Least Skill-Intensive Technology	0.86
$\bar{\gamma}$	Most Skill-Intensive Technology	1.14
ψ	Substitution Across Varieties	0.18

The model has five free parameters (ρ, η,$\bar{\gamma}$, $\bar{Q}$, and ψ) to match eight moments (average Mincerian returns, within-U.S. schooling variation, schooling levels for five quintiles, and elasticity of substitution between U.S. college and high school graduates). Since the model is overidentified, I minimize a loss function over the sum of percentage deviations squared. Table 2 presents the full set of baseline calibrated parameters. The rate of time preference here is higher than the typical $\rho = 0.05$. The elasticity of schooling across countries is slightly larger than the $\eta = 0.4$ used in Bils and Klenow (2000) without education quality, but the identification is very different. The variation in skill intensity across industries is modest. Goods of different skill intensities are very poor substitutes; a value of 2-5 is common for varieties as they are measured in the trade literature, for instance. The model's income predictions are very insensitive to the value of ψ; setting $\psi = 5$ produces virtually identical income differences across countries.

4.2 Model Fit and Income Differences

The model fits the world average Mincerian returns and the within-U.S. variation in schooling quite well. The model predicts within U.S. variation in education levels of 1.82, almost exactly equal to the data; it predicts a world-average Mincerian return to education of 9.13%, in line with the data moment of 9.1%. Figure 9a shows that the predicted Mincerian returns are nearly constant across countries. The lack of trend in the model fits well with the data, although the model is unable to explain any of the observed variation. The elasticity of substitution between high school and college graduates in the U.S. is 1.41, nearly identical to the Katz and Murphy value.



(a) Predicted vs. Actual M_j (b) Predicted vs. Actual S_j

Figure 9: Calibrated Model vs. Data

Figure 9b plots the model's predicted schooling level and the actual schooling level on the left axis, with the U.S. included for reference. The model hits the general upward trend in schooling, but it has some difficulty in matching the exact levels. Figure 9b also plots the quality input on the right axis. The predicted level of schooling tracks quality closely, so the inability of the model to fit the schooling data better is directly due to the quality input. In particular, the model has difficulty with the fact that measured quality shows little variation across the first three income quintiles while average years of schooling doubles.

Now that the model quantitatively replicates the key facts identified in the introduction, I back out the implied income differences. Table 3 gives two different possible ways of breaking down income differences. The model suggests that schooling differences account for a factor of 5.9 of the observed income differences between the United States and the

Table 3: Income Differences Due to Schooling

	U.S.-Poorest	Top-Bottom Quintiles
My Data	70	18.2
Baseline Model	5.9	2.67
Comparison Paper	HJ99[a]	EKR06[b]
Their Data	33	20
Their Model	2.5	4

[a] Hall and Jones (1999), Table 1.

[b] Erosa, Koreshkova, and Restuccia (2006), Table 4. Their calculations are for two economies assumed to have an income ratio of 20, which corresponds well to the actual difference between the top and bottom quintiles.

poorest country. The effect is much larger than was found in the literature that focuses on years of schooling, particularly Hall and Jones (1999).[13] The model suggests that schooling accounts for a factor of 2.67 of the observed income differences between the top 20% and the bottom 20% of countries. The effect here is smaller than predicted in the endogenous school quality literature (Erosa, Koreshkova, and Restuccia 2006).

I use the human capital production function to measure the relative importance of quality and quantity of schooling in the calibrated model. The poorest quintile would be 51% richer if it had the education quality of the United States but its own schooling attainment. It would be 43% richer if it had the same education quality but U.S. schooling attainment. The model suggests education quality differences are more important than education quantity differences.

4.3 Calibration Sensitivity

It is useful to see how sensitive these income predictions are to changes in the main parameters. I focus on the most relevant parameters: ρ, g, σ, $\bar{Q}$, and η. Several of the most important relationships can be closely approximated as a function of parameters and

[13] I compare U.S. to poorest, but the model actually predicts that a few European countries should be richer due to high measured education quality.

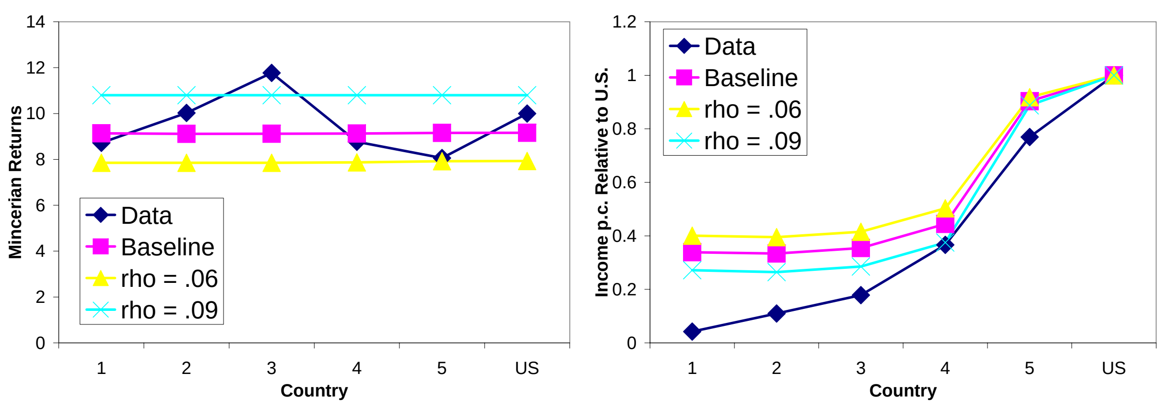
observables:

$$M_j \approx \rho + g\left(\frac{1}{\sigma} - 1\right) \tag{17}$$

$$\frac{S_i}{S_j} \approx \left(\frac{Q_i}{Q_j}\right)^{\eta/(1-\eta)} \tag{18}$$

$$\frac{Y_i}{Y_j} \approx \exp\left(\frac{M}{\eta}(S_i - S_j)\right)\frac{T_j}{T_j - S_j}\frac{T_i - S_i}{T_i} \tag{19}$$

These equations give a feel for how the model reacts to changes in its key parameters. First, increasing ρ has the same effects as increasing g or lowering σ. Equation (17) shows that a higher ρ increases the Mincerian returns to schooling. Since the Mincerian returns are the private value of a year of schooling in the baseline model, equation (19) shows that increasing ρ also increases the model's predicted income variation. Figure 10 shows the results of fixing ρ at alternative values and recalibrating the rest of the model parameters. While a higher ρ would lead to larger income predictions, it also leads to counterfactually higher Mincerian returns.

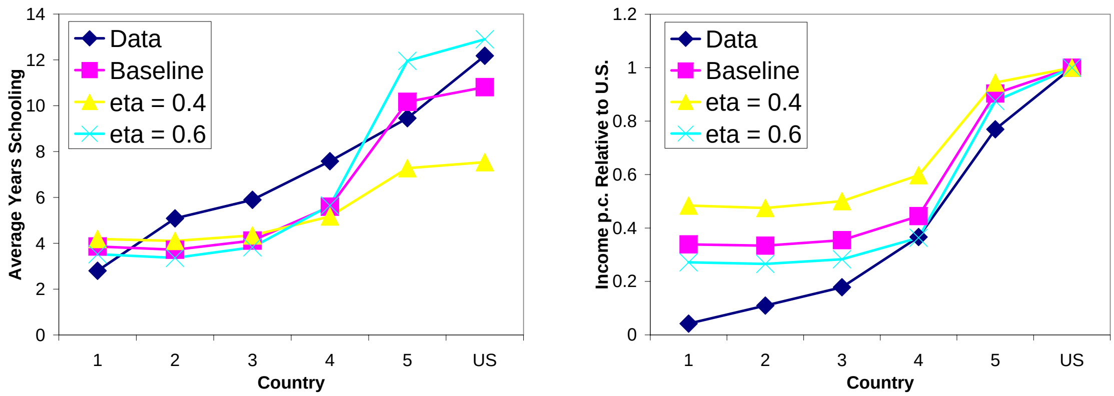


(a) Effects on Returns

(b) Effects on Income

Figure 10: Results of Alternative ρ

Likewise, lowering η has two effects. From (18) it decreases the cross-country variation in schooling predicted by the model. The effects on income predictions are ambiguous: lower η causes less schooling variation, but η also appears directly in (19). Figure 11 shows the effects of fixing η at higher or lower values and recalibrating the model. A higher η leads to larger predicted income variation, primarily through the channel of larger schooling differences.



(a) Effects on Schooling

(b) Effects on Income

Figure 11: Results of Alternative η

Although it does not show up in these equations, $\bar{Q}$ has an obvious role in setting the levels of schooling across countries. Choosing a lower quality normalization leads to lower predicted levels of schooling for all countries. The model also predicts less income variation since there is less schooling variation. Figures for this straightforward case are omitted.

Overall, the model is reasonably robust to changes in the main parameters. It never predicts an income difference of less than a factor of 2 between the richest and poorest 20% of countries for the changes considered. Further, the moments are tightly pinned down by the calibration process, so that changing their value causes the model to produce counterfactual implications elsewhere.

5 A Model with Heterogeneous Workers and Immigrant Selection

The baseline model assumes that workers are all ex-ante identical within a country, with schooling differences arising ex-post only because workers are employed in different industries. Here, I introduce worker heterogeneity into the model. Worker heterogeneity makes it possible to discuss the difference between observed Mincerian returns and the private return to schooling. Additionally, it opens the door to the possibility that immigrants may be different than non-migrants. I show how both of these issues appear in the model and measure their effects on its income predictions.

5.1 Changes to the Model

I incorporate worker heterogeneity into the model through the human capital production function. A worker in dynasty d with $S_j(d,t)$ years of schooling now has human capital given by his human capital production function:

$$H_j(d,t) = \exp\left[\frac{(S_j(d,t)Q_j(d))^\eta}{\eta}\right] \tag{20}$$

The rate of human capital formation per year of schooling is idiosyncratic to the dynasty.[14] It is convenient to disaggregate this term into $Q_j(d) = \bar{Q}_j(1+\tilde{Q}_j(d))$, where $\bar{Q}_j$ is the average school quality in j, and $\tilde{Q}_j(d)$ is the dynasty's mean zero, nonnegative idiosyncratic component. $\tilde{Q}_j(d)$ can represent local school quality variation or heterogeneity in worker ability.

The model has the same basic structure in terms of demographics, preferences, and industries. One change in notation is in order; since workers have different levels of education quality, different wages can be offered to two workers with the same schooling attainment. Hence, it is necessary to write the generalized wage schedule $W_j(S_j, Q_j, \gamma, t)$. It is now convenient to represent the worker's income maximization problem as having two choice variables: the level of schooling, and the industry of employment. That is, workers choose how long to go to school, and when they graduate, they choose the industry that offers them the highest wage. Their problem is still to maximize lifetime income:

$$\omega_j(d,t) = \max_{\gamma_j(d,t),S_j(d,t)} \int_{\tau+S_j(d,t)}^{\tau+T_j} e^{-\int_\tau^t R_j(\theta)d\theta} W_j(S_j(d,t), Q_j(d), \gamma(d,t), t)dt \tag{21}$$

The worker's income maximization problem has two first-order conditions, one for the

[14] It is possible to incorporate worker heterogeneity in other ways. One form common in the labor literature is to generalize the discount rate $\rho(d)$, which would have little effect here. It is also possible to introduce ability as a term that shows up symmetrically to worker-specific TFP, which would leave schooling decisions unchanged but lead to earnings heterogeneity conditional on schooling.

choice of school, and one for the choice of the industry of employment:

$$\int_{\tau+S_j(d,t)}^{\tau+T_j} e^{-\int_\tau^t R_j(\theta)d\theta} \frac{\partial W_j(S_j(d,t), Q_j(d), \gamma(d,t), t)}{\partial S_j(d,t)} dt = e^{-\int_\tau^{\tau+S_j(d,t)} R_j(\theta)d\theta} W(S_j(d,t), Q_j(d), \gamma(d,t), S_j(d,t)+\tau) \quad (22)$$

$$\frac{\partial W_j(S_j(d,t), Q_j(d), \gamma, t)}{\partial \gamma} = 0 \quad (23)$$

The first equation is similar to one in the model with identical workers. Implicitly, the second equation held before, but it is useful to be more explicit about it here.

5.2 Balanced Growth Path Adjustments

The balanced growth path of this model looks somewhat different from the model with identical workers. Again, for exposition I make the assumptions that $\sigma = 1$ and that $T_j - S_j$ large. Workers still go to school until the private return equals their internal discount rate:

$$\frac{\partial \log(W_j(S, Q, \gamma))}{\partial S} = R - g = \rho$$

The optimal schooling level for each worker is:

$$S_j(\gamma) = \left[\frac{\gamma_j(d)\,(Q_j(d))^\eta}{\rho}\right]^{1/(1-\eta)} \quad (24)$$

The heterogeneous quality of workers is the underlying driving force in explaining the cross-section of wages and schooling levels. Mincerian returns are no longer the private return to schooling. Instead, Mincerian returns can be thought of as the change in log-wages across the quality distribution, divided by the change in schooling across the quality distribution:

$$M_j = \frac{\frac{d\log(W_j(S(Q_j(d)), Q_j(d), \gamma(Q_j(d))))}{dQ_j(d)}}{\frac{dS(Q_j(d))}{dQ_j(d)}} = \frac{\frac{\partial \log(W_j(d))}{\partial S_j(d)} \frac{dS(Q_j(d))}{dQ_j(d)} + \frac{\partial \log(W_j(d))}{\partial \gamma_j(d)} \frac{d\gamma(Q_j(d))}{dQ_j(d)} + \frac{\partial \log(W_j(d))}{\partial Q_j(d)}}{\frac{dS(Q_j(d))}{dQ_j(d)}}$$

The model accommodates three channels through which higher quality workers earn higher wages: directly through the quality channel; indirectly through accumulation of higher levels of schooling; and indirectly through working in more skill-intensive industries. Using the workers' first-order condition for industry choice (23) and the assumptions from above, it is possible to evaluate this expression as:

$$M_j = \frac{\rho}{\eta}$$

The Mincerian return is the private return to schooling ρ plus a term measuring the higher wages that are caused indirectly by high quality types going to school longer $\frac{1-\eta}{\eta}\rho$. The latter term is similar to the signalling or sorting effect commonly discussed in the labor literature.

The returns to schooling of immigrants can be evaluated as in Section 3:

$$M_j^i \approx \frac{Q_i(d)}{Q_j(d)} M_j = \left(\frac{\bar{Q}_i}{\bar{Q}_j}\right)\left(\frac{1+\tilde{Q}_i(d)}{1+\tilde{Q}_j(d)}\right) M_j \tag{25}$$

The returns to schooling of immigrants are equal to the education quality of immigrants relative to the education quality of natives with the same education level. For instance, the average Portuguese immigrant in the United States has 7 years of schooling. Then the measured quality is the quality of a Portuguese with 7 years of schooling relative to the quality of an American with 7 years of schooling. There are two potential selection biases: the Portuguese with 7 years of schooling may be atypical (non-migrants have 2 years of schooling), but so may the American (the average American has 12). In the next section I show how to back out the implied selection effects.

The model with heterogeneity interprets the three facts of the introduction differently than the baseline model:

1. Higher education quality causes higher education quantity both within and across countries.

2. Observed Mincerian returns are a mixture of the true benefit of schooling and the effects of high-quality workers sorting into higher schooling levels.

3. The returns to schooling of immigrants are a direct measure of the quality of education of those immigrants relative to the quality of education of a native worker with the same level of schooling. These returns could be biased if either the immigrant or the

American he is compared to is not representative.

In the next section I calibrate a model with these interpretations.

6 Calibration and Selection

6.1 Immigrant Selection

In the simplest selection problem, immigrants are selected on unobserved ability that shows up as a constant in the log-wage regression (Borjas 1987, Borjas 1999). Failing to account for this term is equivalent to an omitted variable problem and leads to biased estimates of the parameters of interest. Bratsberg and Terrell (2002) recognize this problem and use a standard Heckman correction in their paper. They find that it makes little difference to their results.

The model in the previous section suggests that there may be a more difficult problem: workers may be selected on the returns themselves. Then the observed returns of immigrants are correctly measured, but are not representative of non-migrants. One simple example would be if future immigrants attend private schools where education quality is high, while non-migrants attend lower quality public schools; then the measured returns would be the quality of the private schools. The possibility that non-random migration may bias estimates of education based on returns is one of the principle points raised in the Heckman, Layne-Farrar, and Todd (1996) criticism of Card and Krueger (1992).

U.S. immigration policy explicitly selects based on skills, including schooling. In this model, it is high-quality immigrants who have higher schooling. Selecting on schooling is then equivalent to selecting on education quality. To measure the degree of selection, I compare the average years of foreign schooling of immigrants to non-migrants. Data on average years of schooling is available from Barro and Lee (2001) for 78 of 130 countries in the sample in Section 2. For every country except Mexico, immigrants have more schooling than non-migrants. The most extreme case is Sierra Leone, where immigrants have 13.2 years of schooling and non-migrants have 1.65 years, but the average amount of selection by this criteria is six years. I take these differences as evidence of large selection effects.[15]

[15] There is also a slight discontinuity from the fact that the Section 2 data measures average schooling among workers, while the Barro-Lee data is average schooling in the population over 25. Hence, the average American in my sample has 13.5 years of schooling, while Barro-Lee report an American average of 12.2, indicating that Americans are "selected" by 1.3 years. Still, only Mexican immigrants are less selected.

I measure the implied size of the selection effects. To do this, I find the average quality for immigrants and non-migrants from each country that fits perfectly the observed schooling of both groups, as well as the immigrant returns to schooling for each country. The calibrated model measures the implied selection effects that fit cross-country schooling patterns exactly. Fitting all three observations exactly requires consistency with two conditions, both derived from equation (24). The first relates the schooling of immigrants and non-migrants to differences between them:

$$\begin{aligned} \frac{S^j_{US}}{S_j} &= \left(\frac{Q^j_{US}}{Q_j}\right)^{\eta/(1-\eta)} \left(\frac{\gamma^j_{US}}{\gamma_j}\right)^{1/(1-\eta)} \\ &= \left(1+\tilde{Q}^j_{US}\right)^{\eta/(1-\eta)} (\gamma^j_{US})^{1/(1-\eta)} \end{aligned} \tag{26}$$

where the second line comes from using the fact that $\gamma_j \approx 1$ for all countries in the calibrated model. The second relates the average immigrant to an American with the same level of schooling:

$$1 = \left(\frac{\bar{Q}_j(1+\tilde{Q}^j_{US})}{\bar{Q}_{US}(1+\tilde{Q}^{COM}_{US})}\right)^{\eta/(1-\eta)} \left(\frac{\gamma^j_{US}}{\gamma^{COM}_{US}}\right)^{1/(1-\eta)} \tag{27}$$

where I use the superscript COM to denote the appropriate comparison group, Americans with S^j_{US} years of schooling.

Combining these equations yields an expression for country j average quality which depends on the U.S. average quality and terms which account for the two comparisons made between them:

$$\bar{Q}_j = \bar{Q}_{US}\left(\frac{S_j}{S^j_{US}}\right)^{(1-\eta)/\eta} \left((1+\tilde{Q}^{COM}_{US})^{\eta}\gamma^{COM}_{US}\right)^{1/\eta} \tag{28}$$

Equation (28) suggests a way to measure the model implied true education quality of other countries. The equation is structural because the implied values and magnitude of selection effects depend on the calibrated parameter η and the calibrated type of the U.S. control group, $(1+\tilde{Q}^{COM}_{US})^{\eta}\gamma^{COM}_{US}$. In the next section, I show how to use this equation to correct the quality inputs for selection effects and measure the implied selection.

6.2 Calibration

The new calibration differs in three respects from the one done in Section 4. First, observed Mincerian returns now overstate the true return to schooling. Before, the world average Mincerian return of 9.1% was interpreted as the private return, but now it is interpreted as a portion ρ that is private return, plus a portion $\frac{1-\eta}{\eta}\rho$ that is sorting effect. Since the model estimates a lower private return to schooling, it predicts that schooling accounts for a smaller share of observed income differences.

Second, the model introduces two possible sources of variation in schooling attainments within a country: γ and $(1+\tilde{Q}_j(d))$. Schooling attainment and income predictions are determined by a combination of these variables: $(1+\tilde{Q}_j(d))^\eta(\gamma_j(d))$. Hence, instead of choosing the end points for a distribution $[\underline{\gamma}, \bar{\gamma}]$, I choose the end points for a joint distribution $[(1+\underline{\tilde{Q}}_j(d))^\eta \underline{\gamma}_j(d), (1+\bar{\tilde{Q}}_j(d))^\eta \bar{\gamma}_j(d)]$. No attempt is made to disentangle the relative effects of idiosyncratic quality and industry skill intensity.

Finally, I use the equations suggested above to correct the measures of education quality for immigrant selection effects. The equations require an iterative procedure:

1. Take an initial calibration of the model, including η and a distribution $[(1+\underline{\tilde{Q}}_j(d))^\eta \underline{\gamma}_j(d), (1+\bar{\tilde{Q}}_j(d))^\eta \bar{\gamma}_j(d)]$.

2. For each source country j, find the implied $(1+\tilde{Q}_{US}^{COM})^\eta \gamma_{US}^{COM}$ such that a U.S. worker would have the same schooling attainment as the average immigrant.

3. Plug $(1+\tilde{Q}_{US}^{COM})^\eta \gamma_{US}^{COM}$ into equation (28) and evaluate it at the current calibrated η.

4. Update the $\bar{Q}_j$ to reflect the selection and recalibrate the model, implying a new η and distribution $[(1+\underline{\tilde{Q}}_j(d))^\eta \underline{\gamma}_j(d), (1+\bar{\tilde{Q}}_j(d))^\eta \bar{\gamma}_j(d)]$. Return to Step 2.

I construct a new data set of the 79 countries for which I also have source country schooling quality from the Barro-Lee data. I run this iterative procedure over all 79 countries until the model converges on values for η and $[(1+\underline{\tilde{Q}}_j(d))^\eta \underline{\gamma}_j(d), (1+\bar{\tilde{Q}}_j(d))^\eta \bar{\gamma}_j(d)]$. As noted above, Americans in the labor market have more school than the average American; for consistency, I also apply the selection criteria to them, although it makes little difference. Table 4 presents the re-calibrated parameters; those set to outside evidence are the same as in the previous calibration. ρ is now smaller than the standard value of 0.05. η and the end points of the distribution are very similar to before.

Table 4: Alternative Model Calibrated Parameters

Parameter	Role	Value
Calibrated to Fit Data		
ρ	Time Discount Rate	0.034
η	$\varepsilon_{H,S}$	0.49
$\bar{Q}$	Quality Level	0.34
$\underline{\gamma}(\underline{Q}_j(d))^\eta$	Least Skill-Intensive Match	0.83
$\bar{\gamma}(\bar{Q}_j(d))^\eta$	Most Skill-Intensive Match	1.17
ψ	Substitution Across Varieties	0.85

6.3 Results

Correcting for immigrant selection makes a large difference in the quality measures used in the calibration. Table 6 gives the final corrected values for school quality for each of the 79 countries for which it is available. Figure 12 plots the final selection-adjusted quality measures against the measured returns to schooling of immigrants, using the 3 letter ISO country code for each country. The 45 degree line in the figure corresponds to countries for which the corrected and measured values are the same, indicating no selection effects. Countries below the 45 degree line have lower corrected than measured values, indicating higher selection; those above it have higher corrected than measured values, indicating lower selection. Since $\bar{Q}$ is a free parameter in the calibration, it is not possible to sign whether selection is positive or negative. The identity of the countries fits well with expected results. African and Middle Eastern countries are the most selected, while Latin American countries are the least selected.

As expected, the model predicts that schooling accounts for a smaller fraction of observed income differences. Table 5 presents the income comparisons from the baseline model plus those from the model with heterogeneity. The model now makes predictions only slightly larger than in the literature that does not account for quality. The entirety of this difference comes from the fact that in the calibrated model the Mincerian return is broken down as 5.5% true return and 3.5% selection effect. If the IV literature reviewed in Card (1999) is correct and the true value to a year of schooling is about 10% even after correcting for selection effects, then the income predictions presented in Section 5 are more appropriate.

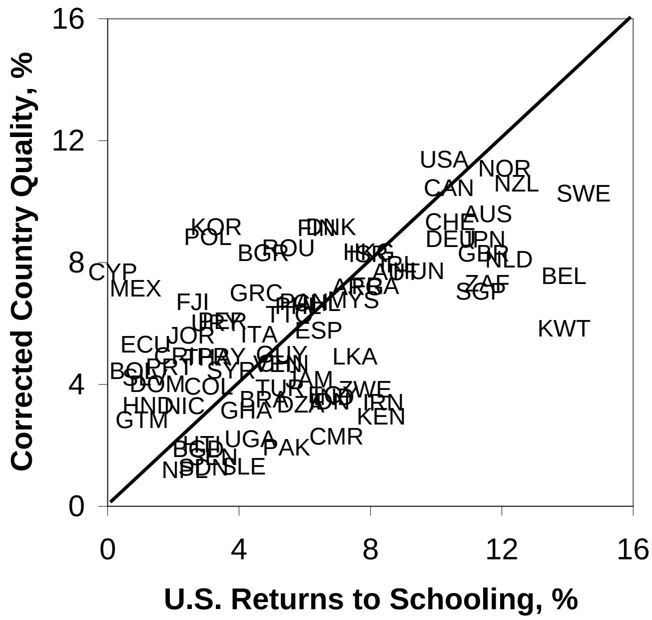


Figure 12: Corrected School Quality vs. Immigrant Returns

7 Conclusion

This paper began by documenting three facts about how a country's education quality interacts with other labor market observables: it is positively correlated with the country's school quantity, uncorrelated with the country's returns to schooling, and positively correlated with the returns to schooling of that country's immigrants to the U.S. I develop a model which fits these facts, and which also suggests using the labor market returns of immigrants to the U.S. as a measure of quality. The implied education quality varies greatly across countries. I use the quality measure as an input to the model and calibrate it to fit the data. The model suggests that quality-adjusted schooling differences account for a factor of 5.9 of the observed income difference between the richest and poorest countries, and a factor of 2.67 of the observed income difference between the richest 20% and the poorest 20% of countries.

I extend the model to include heterogeneity in worker ability. The model allows for immigrant selection, and makes reasonable inferences about the patterns of selection. It also allows for the possibility that observed Mincerian returns overstate private returns to schooling. In this case, schooling differences are less important in accounting for observed income differences. However, there is little evidence in the labor literature that the observed

Table 5: Income Differences Due to Schooling

	U.S.-Poorest	Top-Bottom Quintiles
My Data	70	18.2
Baseline Model	7.5	2.67
Alternative Model[a]	2.87	1.87
Comparison Paper	HJ99[b]	EKR06[c]
Their Data	33	20
Their Model	2.5	4

[a] The baseline model has ex-ante identical workers in each country, while the alternative model has ex-ante heterogeneous workers.

[b] Hall and Jones (1999), Table 1.

[c] Erosa, Koreshkova, and Restuccia (2006), Table 4. Their calculations are for two economies assumed to have an income ratio of 20, which corresponds well to the actual difference between the top and bottom quintiles.

returns exceed the private returns by such an amount.

References

Banerjee, A. V., and E. Duflo (2005): "Growth Theory Through the Lens of Development Economics," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, vol. 1A, chap. 7, pp. 473–554. Elsevier Science, North-Holland Publishers.

Barro, R. J. (1974): "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 82(6), 1095–1117.

Barro, R. J., and J.-W. Lee (1996): "International Measures of Schooling Years and Schooling Quality," *The American Economic Review*, 86(2), 218–223.

——— (2001): "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53(3), 541–563.

Barro, R. J., and X. Sala-i Martin (1999): *Economic Growth*. The MIT Press.

Beaton, A. E., I. V. Mullis, M. O. Martin, E. J. Gonzalez, D. L. Kelly, and T. A. Smith (1996): *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Becker, G. S. (1964): *Human Capital*. University of Chicago Press, Chicago.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy*, 75(4), 352–265.

Bennel, P. (1996): "Rates of Return to Education: Does the Conventional Pattern Prevail in Sub-Saharan Africa?," *World Development*, 24(1), 183–199.

Bils, M., and P. J. Klenow (2000): "Does Schooling Cause Growth?," *The American Economic Review*, 90(5), 1160–1183.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 77(4), 531–553.

——— (1999): "The Economic Analysis of Immigration," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, pp. 1697–1760. Elsevier Science, North-Holland Publishers.

Bratsberg, B., and D. Terrell (2002): "School Quality and Returns to Education of U.S. Immigrants," *Economic Inquiry*, 40(2), 177–198.

Card, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, pp. 1801–1863. Elsevier Science, North-Holland Publishers.

Card, D., and A. B. Krueger (1992): "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100(1), 1–40.

Case, A., and A. Deaton (1999): "School Inputs and Educational Outcomes in South Africa," *The Quarterly Journal of Economics*, 114(3), 1047–1084.

Cervellati, M., and U. Sunde (2005): "Human Capital Formation, Life Expectancy, and the Process of Development," *The American Economic Review*, 95(955), 1653–1665.

Cooley, T. F., and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, chap. 1, pp. 1–38. Princeton University Press.

Erosa, A., T. Koreshkova, and D. Restuccia (2006): "On the Aggregate and Distributional Implications of Productivity Differences across Countries," Working Paper.

Gonzalez, P., C. Calsyn, L. Jocelyn, K. Mak, D. Kastberg, S. Arafeh, T. Williams, and W. Tsen (2000): *Pursuing Excellence: Comparisons of International Eighth-Grade Mathematics and Science Achievement from a U.S. Perspective, 1995 and 1999, NCES 2001028*. U.S. Department of Education, National Center for Education Statistics, Washington, DC: U.S. Government Printing Office.

Gonzalez, P., J. C. Guzman, L. Partelow, E. Pahlke, L. Jocely, D. Kastberg, and T. Williams (2004): *Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003*. U.S. Department of Education, National Center for Education Statistics, Washington, DC: U.S. Government Printing Office.

Griliches, Z. (1977): "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45(1), 1–22.

HALL, R. E., AND C. I. JONES (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *The Quarterly Journal of Economics*, 114(1), 83–116.

HANUSHEK, E. A. (1995): "Interpreting Recent Research on Schooling in Developing Countries," *The World Bank Research Observer*, 10(2), 227–246.

——— (2002): "Publicly Provided Education," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 4, chap. 30, pp. 2045–2141. Elsevier Science, North-Holland Publishers.

HANUSHEK, E. A., AND D. D. KIMKO (1995): "Schooling, Labor-Force Quality, and Economic Growth," NBER Working Paper 5399.

——— (2000): "Schooling, Labor-Force Quality, and the Growth of Nations," *The American Economic Review*, 90(5), 1184–1208.

HANUSHEK, E. A., V. LAVY, AND K. HITOMI (2006): "Do Students Care About School Quality? Determinants of Dropout Behavior in Developing Countries," NBER Working Paper 12737.

HECKMAN, J., A. LAYNE-FARRAR, AND P. TODD (1996): "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings," *The Review of Economics and Statistics*, 78(4), 562–610.

HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2006): "Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1. Elsevier Science, North-Holland Publishers.

HENDRICKS, L. (2002): "How Important Is Human Capital for Development? Evidence from Immigrant Earnings," *The American Economic Review*, 92(1), 198–219.

KALEMLI-OZCAN, S., H. E. RYDER, AND D. N. WEIL (2000): "Mortality decline, human capital investment, and economic growth," *Journal of Development Economics*, 62, 1–23.

KATZ, L. F., AND K. M. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

Klenow, P. J., and A. Rodriguez-Clare (1997): "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," in *NBER Macroeconomics Annual 1997*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 83–103. The MIT Press.

Mankiw, G. N., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2), 407–437.

Manuelli, R. E., and A. Seshadri (2005): "Human Capital and the Wealth of Nations," Working Paper.

Mincer, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

——— (1974): *Schooling, Experience, and Earnings*. Columbia University Press, New York.

Organization for Economic Development (2001): "Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000," OECD Publishing.

——— (2003): "Literacy Skills for the World of Tomorrow: Further Results from PISA 2000. Addendum with Data for Romania," OECD Publishing.

——— (2004): "Learning for Tomorrow's World: First Results From Pisa 2003," OECD Publishing.

Psacharopoulos, G. (1994): "Returns to Investment in Education: A Global Update," *World Development*, 22, 1325–1343.

Psacharopoulos, G., and H. A. Patrinos (2004): "Returns to Investment in Education: A Further Update," *Education Economics*, 12(2), 111–134.

Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander (2004): "Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]," Minneapolis, MN: Minnesota Population Center [producer and distributor], http://www.ipums.org.

Soares, R. R. (2005): "Mortality Reductions, Educational Attainment, and Fertility Choice," *The American Economic Review*, 95(3), 580–601.

Trostel, P., I. Walker, and P. Woolley (2002): "Estimates of the economic return to schooling for 28 countries," *Labour Economics*, 9(1), 1–16.

United Nations Development Programme (2005): *Human Development Report 2005*. United Nations Development Programme.

World Bank (2006): *World Development Indicators*.

A International Test Scores

There are three different test score measures used in the text. The Hanushek-Kimko measure is taken directly from their paper; I am using the QL2* series from Table C1 (Hanushek and Kimko 2000). Construction of the series is explained in greater detail in their working paper (Hanushek and Kimko 1995).

The PISA exam covers three subjects and was administered in two years to different sets of countries. To create an aggregate test score index, I average across subjects. For countries which participated in only one year, I take the data from that year; for countries which participated in both, I take the later data. The TIMSS exam covers two subjects and was administered to different countries in three years. Again, I average across subjects and take the latest year for which data are available. I give preference to the exams administered to older students because more countries participate in these exams. The exact sources for the data are: TIMSS 1995 and 1999, Gonzalez, Calsyn, Jocelyn, Mak, Kastberg, Arafeh, Williams, and Tsen (2000), Table A5.1; TIMSS 2003, Gonzalez, Guzman, Partelow, Pahlke, Jocely, Kastberg, and Williams (2004), Table C2; PISA 2000, Organization for Economic Development (2001), Figures 2.4, 3.2, and 3.5, and Organization for Economic Development (2003) Tables 2.3a, 3.1, and 3.2; PISA 2003, Organization for Economic Development (2004), Figures 2.16b, 6.3, and 6.10.

Both the PISA and TIMSS exams are administered to students who are still enrolled in school. The PISA exam is administered to students between 15 years, 3 months of age and 16 years, 2 months of age. The TIMSS exam is administered to students in the grade which has the highest fraction of 13 year-olds. Beaton, Mullis, Martin, Gonzalez, Kelly, and Smith (1996), Table 2 gives data on the grades actually tested and the amount of school the average student will have had at the time of testing; in general, students from richer countries tend to have fewer years of schooling. Countries were generally allowed to exempt some students from participating in their exams. Gonzalez, Calsyn, Jocelyn, Mak, Kastberg, Arafeh, Williams, and Tsen (2000), Table A2.1 gives an idea of the exclusion rates in the TIMSS study, while Organization for Economic Development (2001), Table A3.1 does the same for the PISA study. Since I focus on these measures as a consistency check for my work, I do not explore the potential biases that may stem from these or other issues of non-comparability. See Organization for Economic Development (2004) and Gonzalez, Guzman, Partelow, Pahlke, Jocely, Kastberg, and Williams (2004) for an explanation of the most recent sampling procedures.

B Model Details

B.1 Final Goods Producer

I assume a zero-profit final-goods producer who purchases $X_j(\gamma, t)$ of each of the intermediate goods and aggregates them into the final good, which she sells to the consumers. The producer uses a CES aggregator with elasticity ψ:

$$Y_j(t) = \left[\int_{\underline{\gamma}}^{\bar{\gamma}} (X_j(\gamma, t))^{1-1/\psi} d\gamma\right]^{\psi/(\psi-1)} \tag{29}$$

I express all time t intermediate good and factor prices in terms of the time t final goods price so that the prices of intermediate goods, capital, and labor are all real. Then the final goods producer's problem is:

$$\max \left[\int_{\underline{\gamma}}^{\bar{\gamma}} (X_j(\gamma, t))^{1-1/\psi} d\gamma\right]^{\psi/(\psi-1)} - \int_{\underline{\gamma}}^{\bar{\gamma}} P_j(\gamma, t) X_j(\gamma, t) d\gamma \tag{30}$$

This problem leads to the standard demand equation:

$$X_j(\gamma, t) = Y_j(t) \left(\frac{1}{P_j(\gamma, t)}\right)^{\psi} \tag{31}$$

B.2 Market Clearing Conditions

There are market clearing conditions in this model for the final goods market, the capital market, the intermediate goods market for each good, and the labor market for each schooling level. These conditions are:

$$Y_j(t) = \int_0^1 \left[c_j(d, t) + \dot{a}_j(d, t) + \delta a_j(d, t)\right] dd \ \forall j, t \tag{32}$$

$$K_j(t) = \int_{\underline{\gamma}}^{\bar{\gamma}} K_j(\gamma, t) d\gamma = \int_0^1 a(d, t) dd \ \forall t \tag{33}$$

$$X_j(\gamma, t) = Y_j(\gamma, t) \ \forall j, t, \gamma \tag{34}$$

$$L_j(\gamma, t) = g_j(S_j(\gamma, t)) \frac{T_j - S_j(\gamma, t)}{T_j} \ \forall j, t, \gamma \tag{35}$$

Here, $K_j(t)$ is the aggregate capital stock, which is equal to the sum of all the asset savings of the workers in the standard way. $g_j(S,t)$ is the density of workers across schooling levels S, with $\int_0^1 g_j(S,t)dS = 1$. Then the first three market clearing conditions are entirely standard. The fourth requires that in equilibrium, the employment of industry γ equals the fraction of workers who have the appropriate level of schooling, $g_j(S_j(\gamma,t))$, times the fraction of those workers who are out of school, $\frac{T_j - S_j(\gamma,t)}{T_j}$.

B.3 Definition of Equilibrium

An equilibrium consists of prices $(R_j(t), W_j(S,t), P_j(\gamma,t))_{t\in[0,\infty)}$ and allocations for intermediate industries
$(K_j(\gamma,t), L_j(\gamma,t), Y_j(\gamma,t), S_j(\gamma,t), H_j(\gamma,t))_{t\in[0,\infty)}$, for the final goods producer $(X_j(\gamma,t), Y_j(t)_{t\in[0,\infty)}$, and for the dynasties $(a_j(d,t), c_j(d,t))_{t\in[0,\infty)}$, as well as a distribution of dynasties across schooling choices $(g_j(S,t))_{t\in[0,\infty)}$, for each country j. These variables must satisfy:

1. The intermediate goods producers' production function, (8) and their FOCs, (9) - (11).

2. The final goods producer's CES production function (29) and its FOC, (31).

3. The dynasties' budget constraints (4) and Euler equations (5).

4. Workers' FOC for schooling, (7).

5. The human capital production function, (6)

6. The market clearing conditions, (32)-(35).

A balanced growth path is an equilibrium such that the variables $R_j, P_j(\gamma), L_j, L_j(\gamma)$ and the distribution $g_j(S)$ are constant, while $W_j, W_j(\gamma), K_j, K_j(\gamma), Y_j, Y_j(\gamma), X_j(\gamma), a_j(\gamma,t)$, and $c_j(\gamma)$ grow at the same rate g as technology. Crucially, the schooling decision is unaffected by the level of aggregate TFP so that schooling decisions are stationary.

C Country Quality Estimates

Table 6: Quality Estimates

Country	Obs	Returns	S.E.	Corrected Quality
Afghanistan	227	6.42	1.01	
Albania	343	-0.95	0.86	
Algeria	82	5.88	1.66	3.39
Antigua-Barbuda	131	10.96	2.05	
Argentina	797	7.62	0.62	7.01
Armenia	325	2.82	0.92	
Australia	464	11.57	1.06	9.2
Austria	153	8.78	1.46	7.46
Azerbaijan	115	8.23	1.75	
Azores	183	0.42	1.16	
Bahamas	108	6.67	2.20	
Bangladesh	602	2.76	0.59	1.97
Barbados	441	3.99	0.96	
Belgium	133	13.89	1.70	7.67
Belize/British Honduras	246	4.82	1.16	
Bermuda	65	9.17	2.36	
Bolivia	371	0.78	0.99	4.44
Bosnia	1135	0.70	0.57	
Brazil	1572	4.76	0.40	3.55
Bulgaria	308	4.74	1.06	8.31
Burma (Myanmar)	332	4.95	0.72	
Byelorussia	314	5.44	1.22	
Cambodia (Kampuchea)	1038	1.31	0.36	
Cameroon	61	6.96	3.06	2.38
Canada	4137	10.39	0.34	9.97
Cape Verde	247	1.99	0.97	
Chile	532	6.37	0.77	6.53
China	8576	5.34	0.13	4.67
Colombia	3667	3.07	0.25	3.96
Costa Rica	475	2.03	0.69	4.91
Croatia	263	3.32	1.02	
Cuba	4842	2.53	0.24	
Cyprus	46	0.16	1.88	7.46
Czech Republic	109	7.94	2.13	
Czechoslovakia	159	5.68	1.55	
Denmark	152	6.80	1.65	8.82

Continued on Next Page

Table 6: Quality Estimates

Country	Obs	Returns	S.E.	Corrected Quality
Dominica	129	2.44	1.76	
Dominican Republic	4343	1.52	0.22	4.05
Ecuador	2121	1.15	0.33	5.26
Egypt/United Arab Rep.	751	6.86	0.77	3.72
El Salvador	7502	1.13	0.17	3.28
Eritrea	127	1.97	1.22	
Ethiopia	493	1.65	0.76	
Fiji	311	2.59	0.99	6.56
Finland	135	6.36	1.54	8.79
France	716	8.15	0.61	7.07
Germany	2736	10.44	0.39	8.46
Ghana	659	4.22	0.76	3.21
Greece	738	4.53	0.59	6.83
Grenada	207	3.61	1.46	
Guatemala	4497	1.04	0.21	2.23
Guyana/British Guiana	1913	5.30	0.41	4.95
Haiti	3707	2.87	0.26	2.11
Honduras	2532	1.23	0.30	3.18
Hong Kong	1177	7.95	0.45	8.08
Hungary	302	9.47	1.09	7.5
India	6378	6.89	0.21	3.63
Indonesia	366	6.75	1.04	3.48
Iran	1480	8.39	0.52	3.45
Iraq	541	1.82	0.53	
Ireland	762	8.84	0.83	7.69
Israel/Palestine	589	7.94	0.70	8.02
Italy	1625	4.61	0.33	5.57
Jamaica	4834	6.14	0.30	4.18
Japan	2269	11.44	0.45	8.44
Jordan	178	2.54	1.30	5.55
Kenya	229	8.33	1.30	3.02
Korea	4549	3.31	0.72	8.82
Kosovo	70	-1.60	1.78	
Kuwait	44	13.90	2.64	5.76
Laos	1475	0.75	0.28	
Latvia	91	5.09	2.36	
Lebanon	468	7.07	0.71	

Continued on Next Page

Table 6: Quality Estimates

Country	Obs	Returns	S.E.	Corrected Quality
Liberia	315	1.14	1.13	
Lithuania	89	7.16	2.46	
Macedonia	143	-0.90	1.33	
Malaysia	325	7.48	0.78	6.62
Mexico	71406	0.85	0.06	5.21
Moldavia	164	4.53	1.52	
Morocco	264	5.74	0.95	
Nepal	81	2.35	1.66	1.27
Netherlands	347	12.22	1.06	7.84
New Zealand	209	12.45	1.58	10.1
Nicaragua	1580	2.34	0.36	3.35
Nigeria	913	4.88	0.62	
Norway	120	12.08	1.90	10.6
Pakistan	1320	5.45	0.41	2.01
Panama	560	5.98	0.88	6.56
Paraguay	68	3.46	2.09	4.89
Peru	2307	3.50	0.38	5.99
Philippines	12283	5.80	0.18	6.43
Poland	3657	3.05	0.31	8.52
Portugal	1531	1.89	0.37	3.7
Puerto Rico	4886	4.68	0.22	
Republic of Georgia	62	8.88	2.55	
Romania	1071	5.50	0.54	8.19
Saudi Arabia	42	0.37	2.77	
Senegal	74	3.22	1.42	1.71
Serbia	81	0.92	1.60	
Sierra Leone	188	4.13	1.47	1.38
Singapore	104	11.36	1.61	6.88
Slovakia	94	8.98	2.56	
Somalia	180	0.42	0.82	
South Africa (Union of)	506	11.55	0.99	7.2
Spain	469	6.42	0.64	5.7
Sri Lanka (Ceylon)	248	7.53	1.10	4.9
St. Kitts-Nevis	98	5.88	2.25	
St. Lucia	118	6.00	1.85	
St. Vincent	159	6.98	1.28	
Sudan	115	2.92	1.44	1.36

Continued on Next Page

Table 6: Quality Estimates

Country	Obs	Returns	S.E.	Corrected Quality
Sweden	240	14.50	1.41	10.1
Switzerland	210	10.43	1.49	9.15
Syria	315	3.76	0.82	4.46
Taiwan	1618	6.88	0.47	
Tanzania	73	15.41	2.52	
Thailand	851	3.05	0.46	4.86
Tonga	111	-0.80	1.60	
Trinidad & Tobago	1567	5.53	0.54	6.18
Turkey	437	5.24	0.72	3.92
Uganda	100	4.35	1.81	2.3
Ukraine	1989	6.73	0.45	
United Kingdom	4449	11.45	0.35	8.02
United States	4122274	10.25	0.01	10.8
Uruguay	188	3.29	1.28	5.91
Uzbekistan	150	4.05	1.75	
Venezuela	536	5.18	0.70	4.65
Vietnam	8276	2.51	0.15	
Western Samoa	86	1.85	1.56	
Yemen Arab Republic (North)	93	2.13	1.12	
Yugoslavia	542	2.66	0.65	
Zimbabwe	78	7.84	2.03	3.87

Note: Country is the country name as it is recorded in the Census files. Obs is the number of observations in the 2000 5% PUMS meeting the sample restrictions. Returns are the log-wage returns to schooling. The returns are measured in percentage points. S.E. is the standard error of the returns. Corrected quality is the immigration selection corrected measure of source-country quality as described in Section 6.