

IZA DP No. 7089

Does Education Matter for Economic Growth?

Michael S. Delgado
Daniel J. Henderson
Christopher F. Parmeter

December 2012

Does Education Matter for Economic Growth?

Michael S. Delgado
Purdue University

Daniel J. Henderson
*University of Alabama
and IZA*

Christopher F. Parmeter
University of Miami

Discussion Paper No. 7089
December 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Does Education Matter for Economic Growth?*

Empirical growth regressions typically include mean years of schooling as a proxy for human capital. However, empirical research often finds that the sign and significance of schooling depends on the sample of observations or the specification of the model. We use a nonparametric local-linear regression estimator and a nonparametric variable relevance test to conduct a rigorous and systematic search for significance of mean years of schooling by examining five of the most comprehensive schooling databases. Contrary to a few recent papers that have identified significant nonlinearities between education and growth, our results suggest that mean years of schooling is not a statistically relevant variable in growth regressions. However, we do find evidence (within a cross-sectional framework), that educational achievement, measured by mean test scores, may provide a more reliable measure of human capital than mean years of schooling.

JEL Classification: C14, J24, I20, O10, O40

Keywords: mean years of schooling, human capital, irrelevant variables, significance testing, nonparametric

Corresponding author:

Daniel J. Henderson
Department of Economics, Finance and Legal Studies
University of Alabama
Tuscaloosa, AL 35487-0224
USA
E-mail: djhender@cba.ua.edu

* We thank the editor Jon Temple, Jenny Minier, Thanasis Stengos, two anonymous referees, and participants at the 2010 Southern Economic Association Conference and the University of Palermo for helpful comments and suggestions.

I. Introduction

There is a large literature that focuses both on the significance of education in an international growth regression and the sensitivity of such regressions to the functional form of the model (Kalaitzidakis *et al.*, 2001). Consistent with this emphasis on functional form sensitivity is the lack of consensus within the literature on whether mean years of schooling is a robustly significant correlate of economic growth. Our goal is to investigate the relevance and statistical significance of mean years of schooling in a generalized empirical growth model that does not require a priori specification of a functional form that may bias the resulting estimates of schooling. Failure to find statistical relevance of schooling in our analysis implies that significance of mean years of schooling in a standard growth model is an artifact of model misspecification.

Much of the existing literature has emphasized the importance of modeling heterogeneity and nonlinearities within the growth process. Temple (2001) and Kalaitzidakis *et al.* (2001), for example, argue that there is a significant nonlinear relationship between economic growth rates and schooling. Durlauf *et al.* (2001) argue that the relationship between growth and initial income is also significantly nonlinear. Sianesi and Van Reenen (2003) provide a thorough review of many econometric issues frequently arising in empirical growth studies, such as parameter heterogeneity, model uncertainty, and nonlinearities. Equally compelling arguments for nonlinearities with respect to schooling come from the micro-foundations of education. For example, Bils and Klenow (2000) argue that human capital is best modeled as a nonlinear function of schooling, and develop a model of human capital and schooling that is based on a typical Mincer model of schooling. Indeed, the earliest microeconomic models of earnings and schooling involved nonlinear relationships (Mincer, 1974). Hence, while the exact relationship between economic growth rates and schooling is not known, it is clear that simple linear econometric models are not sufficient.

One advantage of our fully nonparametric approach is that our model allows for nonlinearities in the growth process for *all* of the variables in our model. This stands in contrast to parametric and semiparametric models that require specification of at least part of the empirical model, often assuming some part of the growth process is linear and homogeneous across countries. The nonparametric growth model also has an advantage over model averaging techniques (Fernández *et al.*, 2001) that are often employed to search for robustness of variables in growth models, because model averaging techniques typically assume that the underlying growth model is linear.¹ While our approach does not allow for simultaneous comparison of numerous potential growth determinants, as does a parametric model averaging approach, the results from our model do not hinge upon correct specification of the functional form of the model.

To further motivate our interest in examining the statistical significance of schooling in a nonparametric growth model, Table 1 provides a brief sample of recent research focusing on the effect of educational attainment on economic growth.² It is clear from the table that a

¹In principle, a model averaging analysis may assume the underlying growth model is nonlinear, but would still typically require specification of the nonlinearities potentially present in the model.

²Educational attainment is traditionally measured using either enrollment rates or mean years of schooling. According to Mankiw *et al.* (1992), which measure of educational attainment should be used depends on ‘whether the available data on human capital correspond more closely to the rate of accumulation (s_h) or to the level of human capital (h).’ See also Gemmell (1996) for a more formal discussion of the difference between the stock and flow of human capital in growth regressions.

consensus fails to exist on the empirical significance of educational attainment in growth regressions. Indeed, there are a variety of parametric, semiparametric, and nonparametric studies that reach widely different conclusions regarding the statistical significance of education in a growth model. The table shows that it is unclear which econometric specification is appropriate in order to correctly capture fundamental aspects of the growth process, and we conclude that the statistical significance of schooling in growth regressions may not be robust across different econometric specifications.

Our nonparametric model is analogous to the benchmark augmented Solow model that includes initial income, investment in physical capital, population growth, and schooling. We include indicator variables to account for region and year fixed effects. We have opted to use a parsimonious specification since the focus here is on the effect of schooling on economic growth and the inclusion of additional variables will most likely introduce collinearity and measurement error into the model. Krueger and Lindahl (2001), for example, note that many other potentially relevant factors of growth may likely depend heavily on education levels (e.g. fertility rates). Moreover, Durlauf *et al.* (2005) have identified over 100 potential growth determinants found in the empirical growth literature. It is not feasible to include all potential growth determinants one might imagine in a single regression, nor would it ever completely remove the possibility of an omitted variable bias. Thus, to simplify issues, and allow a more focused analysis of schooling and economic growth, we include only the ‘traditional’ Solow variables.

It is important in any parametric or nonparametric growth model to recognize the potential endogeneity of schooling. One advantage of using mean years of schooling to proxy for human capital in an empirical growth model is that mean years of schooling is less likely to be correlated with contemporaneous macroeconomic shocks that also effect growth rates, especially over shorter time horizons. Enrollment rates do not have this advantage. Exogenous shocks that affect growth rates could lead to an immediate or rapid change in enrollment rates. However, increased school enrollment, for example, will not lead to higher average schooling of the workforce until the new enrollees have finished their schooling and have entered the workforce. Hence, any shocks to growth rates that cause changes in enrollment rates will not have large effects on mean years of schooling over short time horizons, particularly for schooling measures based on the population 25 years old and above. For this reason, we do not focus on any potential endogeneity of mean years of schooling. Implicit in this approach is the assumption that investment in schooling is not made in anticipation of growth (Bils and Klenow, 2000), or that this effect is not strong enough to influence our results.

We address the issue of data construction for mean years of schooling by considering five different schooling databases that have received considerable attention in the empirical literature (Nehru *et al.*, 1995; de la Fuente and Doménech, 2002; Cohen and Soto, 2007; Lutz *et al.*, 2007; Barro and Lee, 2010). Each database has been created specifically to overcome potential weaknesses and inconsistencies identified in competing databases or previous versions of the same database. Hence, data construction techniques and measurement error have been rigorously addressed across each of the education databases, and by considering each separately we can be certain that our results do not hinge upon the data source or methods by which our schooling data are constructed. See Sianesi and Van Reenen (2003) and Pritchett (2006) for a thorough discussion of the measurement of the stock of education of a country and potential issues concerning measurement error in growth regressions.

We use two nonparametric procedures to investigate the significance of schooling in our

TABLE 1
Sample of recent literature investigating human capital and economic growth

<i>Paper</i>	<i>Human capital data</i>	<i>Method</i>	<i>Summary and significance</i>
Barro (1991)	Primary and secondary school enrollment rates	OLS cross-section	Significant, positive effect of human capital on economic growth
Mankiw <i>et al.</i> (1992)	Secondary school enrollment rates	OLS cross-section	Significant, positive effect of human capital on economic growth
Durlauf and Johnson (1995)	Secondary school enrollment rates	OLS cross-section and regression tree	Human capital significance depends on sample of nations included in regression
Islam (1995)	Barro and Lee (1993)	OLS panel data with dummy variables	Insignificant and negative effect of human capital after controlling for fixed effects
Nonneman and Vanhoudt (1996)	Secondary school enrollment rates	OLS cross-section	Insignificance of human capital after controlling for 'technological know-how'
Liu and Stengos (1999)	Secondary school enrollment rates	OLS and semiparametric partial linear model	Insignificance of human capital in the parametric model; human capital is insignificant and linear in the semiparametric model
Barro (2001)	Barro and Lee (2001)	3SLS with panel data	Significant, positive effect of male secondary education; insignificant effect of female and primary male education
Durlauf <i>et al.</i> (2001)	Secondary school enrollment rates	Semiparametric smooth coefficient model	Significant, nonlinear effect of human capital conditional upon initial income estimates
Kalaitzidakis <i>et al.</i> (2001)	Barro and Lee (1996)	OLS and semiparametric partially linear regression model	Insignificance in parametric models; significant nonlinearities in semiparametric models
Temple (2001)	Barro and Lee (1993, 2000)	OLS and Least Trimmed Squares	Evidence of sensitivity to outliers; nonlinearities in education; tentative evidence of significance of education
Sala-i-Martin <i>et al.</i> (2004)	Barro and Lee (1993)	Bayesian Averaging of Classical Estimates	Significant relationship between growth and primary schooling; insignificant relationship between growth and higher education
Maasoumi <i>et al.</i> (2007)	Barro and Lee (2001)	OLS panel data with dummy variables and local-linear least-squares	Insignificance of human capital in the OLS model and significance in the LLLS model
Minier (2007)	Barro and Lee (2001)	OLS cross-section and regression tree	Positive, significant effect of human capital in baseline regressions; insignificance when controlling for policy and executive constraints
Durlauf <i>et al.</i> (2008)	Barro and Lee (2001)	Bayesian model averaging	Little evidence that human capital is significant and robust
Henderson (2010)	Barro and Lee (2001)	Nonparametric local-linear least-squares	Insignificance for human capital on economic growth

nonparametric growth model. First, we examine the statistical significance of the partial effects of schooling in our nonparametric model to see whether or not a marginal change in mean years of schooling leads to a significant change in growth rates. Second, we employ a nonparametric test of variable relevance (Lavergne and Vuong, 2000) to test the null hypothesis that schooling is irrelevant in our model. We find that our nonparametric regression estimates yield

statistically insignificant partial effects of schooling on growth rates, and our nonparametric variable relevance test fails to reject the null hypothesis of variable irrelevance for nearly all of the specifications we consider. We find relevance for mean years of schooling for two out of five specifications using a nonparametric median regression model, but note that the partial effects for each of these specifications are neither statistically nor economically significant. By contrast, we apply the same nonparametric procedures to our other covariates (initial income, investment, and population growth) and find that each of these variables are generally relevant and statistically significant correlates of economic growth. While there is less uncertainty about the significance of these variables within the empirical growth literature, our results provide evidence that their significance is robust to arbitrary forms of nonlinearities, heterogeneity, or model specification.

Even though nonparametric estimators are local estimators, they are not immune to outliers. It is well documented that nonparametric methods may be sensitive to outliers, especially in the context of a conditional mean model or empirical economic growth models (given the small sample sizes available). We follow Temple (1998, 2000) and investigate the sensitivity of our primary results to outliers. We use a minimum covariance determinant estimator to identify and remove outliers, a nonparametric median regression estimator and testing procedure, and a measurement error corrected education dataset (Portela *et al.*, 2010). We find that the irrelevance of schooling is robust to our outlier sensitivity assessments. We next consider 5-, 10-, and 20-year lagged effects of schooling on growth as well as a variety of different subsamples, including OECD/non-OECD countries and male/female subsamples, and fail to find relevance of schooling.

Therefore, we conclude that mean years of schooling is not a relevant variable in an empirical growth model. Our result is robust to different functional form assumptions that may be imposed on the growth model, to the choice of aggregate schooling database, and the construction methods incorporated into each. Any empirical significance found for education in previous research using just the Solow variables is most likely because of model misspecification. We caution that our results do not necessarily imply that schooling is a poor measure of human capital, only that schooling is not a relevant factor in determining growth rates. Indeed, Pritchett (2001) argues that schooling may not be associated with higher growth rates because (i) educated workers may be motivated to participate in socially unproductive activities - ‘piracy’; (ii) a surplus of skilled labor has suppressed wages and dampened growth; and (iii) that poor quality of schooling has not translated into any increase in human capital. Only in the third case is schooling a poor proxy for human capital. Hence, one implication of our results is that further research into alternative and complementary measures of human capital, such as test scores and health-based measures of human capital, may be warranted.

In order to provide a brief exploration into the potential of alternative human capital measures, we consider a cross-section of the human capital quality data constructed by Hanushek and Kimko (2000) and find variable relevance as well as robust statistical significance of the partial effects of education quality. Bearing in mind the limited capacity of the Hanushek and Kimko (2000) data because of its small sample size, our results suggest that test score quality measures may provide a more reliable measure of human capital than mean years of schooling. Indeed, significant differences in the quality of schooling across countries likely implies that economic growth rates are more highly correlated with educational achievement than attainment (Schoellman, 2012).

The rest of this paper is outlined as follows. Section 2 details the empirical methodologies

used throughout the analysis. Section 3 provides a description of the data. Section 4 presents the results from the baseline sample regressions. Section 5 presents a summary of the results from a variety of robustness checks, which include outlier sensitivity, lagged schooling effects, alternative subsamples and specifications, and measures of educational achievement. Section 6 concludes.

II. Methodology

Our econometric strategy is based on a nonparametric version of the canonical growth regression model. Our model is analogous to the conditional convergence equation in Mankiw *et al.* (1992), in which human capital is measured as a stock instead of a flow. The nonparametric regression model is

$$g_i = m(x_i) + u_i, \quad i = 1, 2, \dots, nT \quad (1)$$

in which g_i is the growth rate of a particular country in a particular time period, $m(\cdot)$ is a smooth function of unknown form, x_i is a p -dimensioned vector of explanatory regressors, and u_i is a mean zero error term. Specifically, we define $m(x_i) \equiv E[g_i|x_i]$; that is, we define $m(x_i)$ to be the conditional mean of g_i given the regressors in x_i . Our regressors include initial income, investment in physical capital, population growth, education (defined as mean years of schooling, to proxy for the stock of human capital), and indicator variables controlling for region and year fixed effects.

We employ a dual nonparametric approach to check for robustness of schooling in our nonparametric growth model. We first employ nonparametric local-linear least-squares to estimate Eq. (1), and look for statistical significance of the partial effects of education on the economic growth rate. This strategy is analogous to using a t-test for coefficient significance in a standard parametric linear regression model, albeit at fixed points. However, because our regression model is nonparametric, we have the added advantage of avoiding parametric model misspecification that may influence whether or not we find statistical significance of the partial effects. As in a parametric t-test of coefficient significance, we interpret statistical insignificance of the partial effects to imply that the regressor (at that point) does not have a significant effect on the economic growth rate, holding constant other variables in the model.

The test of significance for the partial effects of education on economic growth is informative as it tells us whether or not a marginal change in mean years of schooling is correlated with a significant change in economic growth rates. However, this test is limited in its ability to determine whether or not education is a robust variable. Finding that the partial effect(s) of education on economic growth is insignificant does not allow us to exclude education from the model. Hence, we also employ a formal nonparametric test of variable relevance (Lavergne and Vuong, 2000) to help us determine whether or not education is a robust and relevant regressor, and this test provides the backbone of our analysis. If the nonparametric test of variable relevance fails to reject the null hypothesis that education is an irrelevant variable, then we have formal statistical evidence that education, as measured by mean years of schooling, does not belong in a growth model such as Eq. (1). Of course, if the test rejects the null hypothesis that education is an irrelevant variable, then we have robust statistical evidence that education, as measured by mean years of schooling, is a robust correlate of the rate of economic growth.

The local-linear least-squares estimator of Eq. (1) is derived from a first-order Taylor ex-

pansion of Eq. (1) around a point x in the subset of continuous regressors in x_i :

$$g_i \approx m(x) + (x_i^c - x^c)\beta(x) + u_i \quad (2)$$

in which x_i^c denotes the continuous regressors in x_i , and $\beta(x)$ denotes the partial derivative of $m(x)$ with respect to x . The local-linear estimator provides an estimate of $\delta(x) \equiv \begin{pmatrix} m(x) \\ \beta(x) \end{pmatrix}$ via

$$\hat{\delta}(x) = (X'K_h(x)X)^{-1}X'K_h(x)g \quad (3)$$

in which X is a matrix containing a column of ones and the continuous regressors in x_i with the i th row being $X_i = [1 \ (x_i^c - x^c)]$, $K_h(x)$ is a weighting matrix with $k_h(x_i, x)$ on the diagonal, and $k_h(x_i, x)$ is a generalized product kernel function (Racine and Li, 2004) with bandwidth h that admits both continuous and discrete variables.³ We determine optimal bandwidths using least-squares cross-validation, which selects the bandwidth by minimizing the following criterion function

$$\min_h \sum_{i=1}^n [g_i - \hat{m}_{-i}(x_i)]^2 \quad (4)$$

where $\hat{m}_{-i}(x_i)$ is the leave-one-out estimate of $m(x_i)$. Standard errors for the partial effects from the local-linear least-squares estimator are obtained using a wild bootstrap.

The intuition for the local-linear estimator is that $\hat{\delta}(x)$ is obtained from a locally weighted linear regression of g_i on x_i . The generalized product kernel function, $K_h(x)$, serves as the local weighting function, assigning more weight to observations close to x , and less weight to observations further from x . See Maasoumi *et al.* (2007), Henderson (2010), and Henderson *et al.* (2012) for recent examples of applications using the local-linear least-squares estimator in an empirical growth setting.

We use the nonparametric test for variable relevance proposed by Lavergne and Vuong (2000) to determine whether or not human capital, as measured by mean years of schooling, is a variable that belongs in Eq. (1). Specifically, Lavergne and Vuong (2000) consider a test of the null hypothesis

$$H_0 : m(\tilde{x}_i) = m(x_i) \quad (5)$$

in which \tilde{x}_i is a subset of x_i ; in our case, \tilde{x}_i excludes mean years of schooling. That is, we wish to test the null hypothesis that mean years of schooling does not belong in Eq. (1), against the alternative hypothesis

$$H_1 : m(\tilde{x}_i) \neq m(x_i), \quad (6)$$

that mean years of schooling is a relevant variable in Eq. (1). In order to properly test the null hypothesis, Lavergne and Vuong (2000) propose the following test statistic

$$I_n = \frac{(n-2)!}{n!} \sum_{i=1}^n \sum_{j=1}^n \hat{u}_i \hat{f}_{-i}(\tilde{x}_i) \hat{u}_j \hat{f}_{-j}(\tilde{x}_j) k_h(x_i, x_j) \quad (7)$$

in which $\hat{u}_i = g_i - \hat{m}_{-i}(\tilde{x}_i)$ is an estimate of the residuals from the leave-one-out first stage

³Partial effects for discrete variables are obtained separately as a counterfactual change in the estimated regression function as the discrete variables switches from the base value to another. In our application, we focus solely on partial effects of continuous regressors.

restricted regression of g_i on \tilde{x}_i , $\hat{f}_{-i}(\cdot)$ is a leave-one-out estimate of the probability density of the covariates, $f(\cdot)$, and $k_h(x_i, x_j)$ is a product kernel weighting function. Lavergne and Vuong (2000) recommend using a rule of thumb bandwidth for variable relevance testing as it was shown to provide robust size and power in their simulation experiments. We use a wild bootstrap to approximate the distribution of the test statistic under the null hypothesis, and report p-values based on a standardized version of Eq. (7).

III. Data

With the exception of the education data and categorical indicators, our data are derived from the Penn World Table Version 7.0, (Heston *et al.*, 2011). Our education datasets are those constructed by Nehru *et al.* (1995), hereafter NSD; de la Fuente and Doménech (2002), hereafter DD; Cohen and Soto (2007), hereafter CS; Lutz *et al.* (2007), hereafter IIASA-VID; and Barro and Lee (2010), hereafter BL.⁴

Penn World Table 7.0

All international growth rates are based on the Penn World Table Version 7.0 (Heston *et al.*, 2011). The data include measures of per worker GDP growth, initial income, the investment rate, and working-age population growth. Initial income, investment, and population growth are measured in logs. Per worker GDP growth is the average growth rate over each 5-year period, derived from the chain-weighted real GDP per worker series (RGDPWOK). Initial income is measured as the log of real income per worker at the beginning of each 5-year period (again taken from RGDPWOK). The investment rate is the log of the average investment share of income over each period, derived from KI in the Penn World Table. The population growth variable is the log of the average growth rate over each 5-year period derived from the total working-age population (POP). Our Penn World Table 7.0 data series cover the period from 1950-2005, however the time frame varies due to differences in human capital data availability. We note that in cases in which the 5-year period did not contain five complete annual observations, we calculated the growth rates based on the available data.

Categorical variables

A regional indicator was created based on the regional country classifications of the World Bank (available online at www.worldbank.org).⁵ We have chosen to use the World Bank regional classification system instead of other common classifications (e.g., OECD versus non-OECD) since our aim is to capture geographic similarities and other methods have a tendency to mask relevant information or provide economic, not geographic, information. For example, an OECD/non-OECD indicator is largely based on economic similarities and thus will not capture geographic or cultural similarities. Moreover, this will not identify heterogeneity between

⁴The Lutz *et al.* (2007) dataset was constructed as part of a joint research effort conducted by the International Institute for Applied Systems Analysis (IIASA) and the Vienna Institute of Demography (VID).

⁵We also tried using a country level indicator to control for individual country effects but found the smoothing parameter is zero, effectively separating each country into its own individual sample. The regional indicator avoids this complication and effectively controls for regional effects without completely dividing the sample.

any non-OECD nations which can hardly be assumed to be homogeneous. The World Bank's regional classification system seems to provide the most reasonable geographic division.⁶ That being said, the World Bank's 'reference' region is largely made up of OECD nations. At the very least, the regional indicator should still capture substantial geographic heterogeneity among non-OECD nations. We include a single ordered categorical indicator to control for time effects, spanning the interval 1950-2005.

Education data

Nehru et al. (1995)

The NSD dataset was constructed by Nehru *et al.* (1995), although the version we use is from Duffy and Papageorgiou (2000). The database contains annual estimates of mean years of educational attainment for the working-age population, aged 15-64, for a sample of 82 countries from 1960-87.⁷ NSD use enrollment rates via the perpetual inventory method, adjusting for grade repetition as well as dropout and mortality rates, to construct educational attainment estimates. The primary sources used are UNESCO enrollment data and yearbooks and the enrollment data estimated by Mitchell (1982). This paper uses 5-year averages instead of annual estimates as given by NSD. Annual estimates are typically thought to be too sensitive to business cycle fluctuations and may not serve as a reliable proxy for human capital.

de la Fuente and Doménech (2002, 2006)

The DD human capital data come from de la Fuente and Doménech (2002, 2006). The estimates are for average years of total schooling for the population aged 25+ for a sample of 21 OECD countries over 5-year intervals from 1960-95. This is the only human capital database considered in this paper that does not provide estimates for the population aged 15+. The data also include estimates of the share of the population with a given level of educational attainment over various education levels. DD use census and survey data to construct educational attainment estimates and rely on interpolations and projections to fill in missing values. They explicitly avoid the use of enrollment data due to inconsistent time profiles. Primary sources used are OECD and national census/survey databases.

Cohen and Soto (2007)

The CS data measure average years of schooling for the working-age population for two aggregate age groups (15+ and 25+) and for the entire population aged 15+. The database also contains estimates of the share of the population receiving a primary, secondary, or higher education. The sample is estimated over 10-year intervals covering the period 1960-2010 (the period from 2001-10 is a projection) and includes 95 countries. CS use population shares of educational attainment based on census data to construct average years of schooling using enrollment rates, adjusted for mortality rates, to fill in missing observations. Primary sources used

⁶The World Bank's geographic classification system does not include every country in our dataset. For any country not classified by the World Bank we classified it according to the nearest geographic region.

⁷The original NSD database contained human capital data for 87 countries. 5 countries were dropped by Duffy and Papageorgiou (2000) due to the availability of data for other variables used in their analysis.

to estimate educational attainment are the OECD's education database, UNESCO's Statistical Yearbook and Statistics and national agencies' censuses. Since the data are available from Cohen and Soto in 10-year intervals, we linearly interpolated the data to obtain 5-year intervals in order to maintain consistency with alternative human capital databases. Regression results utilizing the original 10-year intervals are later reported as a robustness check.

IIASA-VID (Lutz et al. 2007)

The IIASA-VID database includes estimates of average years of schooling and proportions of educational attainment according to age (5-year age groups) and gender for the working-age population aged 15+ for a sample of 120 countries over the years 1970-2000. The database also includes aggregate measures of mean educational attainment for three aggregate age groups (15+, 25+ and 65+) for males, females and both sexes together. IIASA-VID uses census and survey data to construct human capital stock estimates for a base year (2000) and back projects until 1970 to obtain estimates of human capital based on age and gender, adjusting for mortality and migration rates. The basic idea behind these back projections is that the stock of human capital for each age cohort in the year 2000 is directly related to the stock of human capital for the previous age cohort five years prior, adjusted for mortality. For example, the stock of human capital for men aged 50 in the year 2000 must be the same as the stock of human capital for men aged 45 in the year 1995, after adjusting for mortality rates. This demographic back projection technique allows the researcher to obtain estimates of the human capital stock for different age cohorts across both genders without relying on historical survey or census data. Another benefit of this method is that it avoids issues that arise from changes over time in the methods used for data collection. This is the only human capital dataset to utilize this approach.

Barro and Lee (2010)

The BL dataset is primarily a response to critical observations of older editions, for example Barro and Lee (2001). The dataset extends coverage of mean years of schooling data to 146 countries over 5-year intervals from 1950-2010, disaggregated by gender and age (15+ and 25+). Census and survey data obtained from UNESCO and Eurostat are used to construct estimates of various levels of educational attainment. Missing observations are estimated by extrapolating backwards and forwards from census and survey data. Estimates are corrected for mortality rates that are allowed to differ across different education cohorts. Preliminary testing (Barro and Lee, 2010) shows that these new estimates have successfully corrected the inconsistencies identified in previous BL versions providing smoother time profiles for educational attainment. Their estimated reliability ratios suggest BL is more reliable than CS, which was shown by Cohen and Soto (2007) to be more reliable than Barro and Lee (2001).

Preliminary discussion

As can be seen from the descriptions above, there is still some dispute as to the most suitable methods for estimating mean years of schooling. One theme emerging in the most recent education measures (i.e. DD, CS, IIASA-VID and BL) is a distrust of enrollment rates to serve as a reliable proxy for human capital. de la Fuente and Doménech (2006) point out numerous implausibilities in estimates based entirely on enrollment data. Census and survey data are

believed to provide a more accurate picture of current educational conditions. Currently, how to properly fill in missing observations appears to be the main point of dispute. CS employ the most ‘traditional’ approach by using enrollment rates, whereas DD and IIASA-VID rely solely on projections.

The most commonly-used dataset in growth empirics is Barro and Lee (2001), however recent research has argued that new, alternative education measures are empirically more relevant in explaining cross-country growth variation and tend to be more robust when compared to the Barro and Lee (2001) estimates (Barro and Lee, 2010; Lutz *et al.*, 2007; Cohen and Soto, 2007; de la Fuente and Doménech, 2006). The NSD database, according to Cohen and Soto (2007), has not enjoyed the empirical success of the Barro and Lee (2001) dataset since it fails to make use of census data and in some cases their estimates do not coincide with information available from censuses; see also de la Fuente and Doménech (2006) for a similar critique. Given the release of the newest BL dataset, it is perhaps too soon to tell which of the alternative education datasets will prove most reliable and useful; however Savvides and Stengos (2009) name CS as the most likely competitor of Barro and Lee (2001).⁸ Preliminary testing (de la Fuente and Doménech, 2006; Cohen and Soto, 2007; Lutz *et al.*, 2007) has shown that the DD, CS and IIASA-VID estimates yield qualitatively similar results and are in general jointly qualitatively different from Barro and Lee (2001).

One issue that needs to be addressed is the question of comparability between different education measures. Due to the differences in either time spans or samples of countries, different education datasets will have different sample sizes. This is most notably an issue for the DD dataset. For DD, the results obtained are not necessarily directly comparable to those from alternative, more comprehensive sources since DD does not include non-OECD nations. In light of our goals for this paper, differences between the samples of observations used across different educational databases work to strengthen the generality of our results. If we find that mean years of schooling is irrelevant across each of the different education databases, in spite of the differences between samples of observations, we have further reassurance of the generality of our results than had the subsamples of observations been the same. We use a variety of subsamples of observations with each education database in order to ensure that our results do not depend on the observations in the sample or the education database used to proxy for human capital.

We calculate the sample correlations between educational attainment measures to obtain initial insights into how similar (or different) alternative attainment measures are. We first calculate the sample correlation matrix for all five education datasets for a consistent sample of observations. The sample spans the period from 1970-85 and has a total of 84 observations. We find the highest correlations are between CS and DD (0.93), DD and IIASA-VID (0.84), and CS and IIASA-VID (0.76). The lowest correlations are between IIASA-VID and NSD (0.26) and DD and NSD (0.29). In general, NSD is not highly correlated with any of the alternative datasets. We find, however, that these low correlations disappear once we remove DD from the sample and redefine our consistent sample across the remaining four datasets, yielding a new sample of 256 observations spanning the period 1970-85. The highest correlations are between BL and CS (0.93) and CS and IIASA-VID (0.93), yet the lowest correlation is between IIASA-VID and NSD (0.84). We no longer see such dramatic differences in sample correlations across

⁸Savvides and Stengos note, in particular, that DD and NSD are not adequate human capital sources since they only cover a sample of OECD nations and end in 1987, respectively. They make no mention of the IIASA-VID database.

TABLE 2
Parametric estimates for schooling

<i>Parametric estimator</i>	<i>Educational attainment database</i>				
	<i>BL</i>	<i>CS</i>	<i>DD</i>	<i>IIASA-VID</i>	<i>NSD</i>
Pooled OLS	-0.0017	-0.0008	0.0075	0.0025	-0.0028
	0.0018	0.0017	0.0064	0.0024	0.0024
Fixed Effects	-0.0160	-0.0013	-0.0049	-0.0227	-0.0040
	0.0043	0.0044	0.0313	0.0077	0.0067
Between	0.0073	0.0009	-0.0103	0.0034	0.0009
	0.0028	0.0023	0.0079	0.0034	0.0026
First Differences	-0.0361	-0.0268	0.0627	-0.0623	-0.0116
	0.0126	0.0131	0.0767	0.0183	0.0159

Notes: Table reports coefficient estimates and standard errors for schooling only. Statistically significant coefficients at the 5% level are emphasized in bold.

the remaining datasets. Restricting our sample to 546 observations spanning 1970-2000 for only BL, CS and IIASA-VID as these are the most comprehensive education datasets, we find that each of the three pairwise sample correlations are between 0.92 and 0.94. We conclude that for relatively small samples that are restricted to a subset of developed countries, there are substantial differences in correlation across each of the education datasets, but for relatively large samples that include both developed and developing countries, these datasets are highly correlated.

IV. Baseline sample results

Parametric models

Prior to estimating our nonparametric regressions, we first summarize the results from several standard parametric panel data specifications. Specifically, we consider a simple pooled OLS estimator, the fixed effects (within) estimator, the between estimator, and the first differences estimator. The purpose of estimating parametric regression specifications is to verify that our data provide estimates that are consistent with previous empirical research as well as illustrate the sensitivity of typical growth regressions to the estimation procedure. Table 2 contains estimated coefficients for schooling from each of the parametric specifications.

We find the sign and significance of education to be, in part, dependent on the estimation procedure. In many cases, we see educational attainment is statistically insignificant, or negative and significant (Islam, 1995). In terms of population growth, investment, and initial income, we find that the (unreported) coefficient estimate on each regressor typically takes the expected magnitude and sign, as well as statistical significance.

Nonparametric partial effects

We now turn to our nonparametric regression model estimated using local-linear least-squares. Table 3 presents the estimates of the partial effects along with their corresponding wild bootstrapped standard error in parentheses.⁹ Each nonparametric specification includes unordered and ordered categorical indicators to control for region and time effects. Since nonparametric estimation provides partial effects (and therefore standard errors) for each observation, we summarize the partial effects for each regressor at the 25th, 50th, and 75th percentiles. We emphasize that this exercise is analogous to performing t-tests of parameter significance in a standard linear model at each of the three fixed points, and that general statistical insignificance of the partial effects for any regressor suggests that the regressor is not a statistically significant determinant of growth rates. We highlight statistically significant partial effects at the 5% level in bold.

Before presenting the regression results, we point out that Table 3 shows that there are substantial differences between the sample sizes depending on which human capital dataset is included in the regression. The largest datasets are BL, CS and IIASA-VID, spanning the periods 1950-2005, 1960-2005, and 1970-2000, with 1,376, 921, and 764 observations, respectively. DD spans the interval 1960-95 and contains the smallest number of observations, 161, while the NSD dataset is of moderate size, spanning the interval 1960-85 and containing 478 observations. The number of observations available with each dataset is important for a number of reasons, but most importantly the following. First, the larger human capital datasets contain observations for a larger number of non-OECD nations (recall DD only contains OECD nations), and over a longer time horizon. It is important to recognize that data quality likely diminishes with these larger samples. Second, the nonparametric methods we use to assess the robustness of schooling are more reliable with larger sample sizes - it is important to bear in mind the smaller sample size when interpreting the results from the DD dataset.

Table 3 reveals that initial income is generally negative and statistically significant only at the 25th percentile for each specification. The only exceptions are BL for which initial income is not significant at the reported intervals, and for DD in which initial income is negative and significant at both the 25th and 50th percentiles. Henderson *et al.* (2012) find similar results for initial income in a nonparametric growth setting. We find that physical capital investment is positive with a magnitude that is similar to estimated coefficients in the parametric models, and is generally statistically significant at the 50th and 75th percentiles. Population growth is generally negative and statistically significant across each of the five estimated growth regressions, and is consistent with our a priori expectations. Finally, Table 3 shows that schooling is statistically insignificant at each reported percentile for every education dataset. That is, we do not find any evidence that the partial effects (at the quartiles) of mean years of schooling on the economic growth rate are statistically significant for any of the datasets under consideration. We note that the R^2 , defined as the square of the sample correlation between the observed and fitted growth rates, is generally around 0.6, suggesting that the nonparametric growth model is capable of explaining substantially more of the variation in growth rates compared to the standard linear parametric setup ($R^2 \approx 0.28$), regardless of which education measure is used.

⁹The estimates of the standard errors for each of the models estimated in this paper are obtained via 399 replications of a wild bootstrap. We utilize a wild bootstrap as it is robust to heteroskedasticity that would otherwise render a standard residual bootstrap invalid for statistical inference (Cameron and Trevedi, 2005).

TABLE 3
Summary of partial effects and standard errors from the local-linear regression model for the baseline sample

<i>Educational attainment database</i>															
<i>Variable</i>	<i>BL</i>			<i>CS</i>			<i>DD</i>			<i>IIASA-VID</i>			<i>NSD</i>		
	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>	<i>Q1</i>	<i>Q2</i>	<i>Q3</i>
Initial Income	-0.0088	-0.0040	0.0005	-0.0106	-0.0046	-0.0002	-0.0243	-0.0221	-0.0189	-0.0048	-0.0027	-0.0004	-0.0125	-0.0048	0.0003
	0.0098	0.0028	0.0039	0.0050	0.0042	0.0035	0.0062	0.0085	0.0110	0.0024	0.0025	0.0016	0.0053	0.0030	0.0025
Investment	0.0113	0.0191	0.0271	0.0091	0.0167	0.0223	0.0180	0.0258	0.0286	0.0088	0.0185	0.0268	0.0137	0.0212	0.0281
	0.0081	0.0077	0.0067	0.0081	0.0088	0.0076	0.0090	0.0042	0.0044	0.0115	0.0080	0.0091	0.0066	0.0052	0.0065
Population Growth	-0.0501	-0.0331	-0.0126	-0.0619	-0.0422	-0.0241	-0.0428	-0.0221	-0.0161	-0.0495	-0.0366	-0.0154	-0.0455	-0.0351	-0.0170
	0.0136	0.0151	0.0142	0.0186	0.0061	0.0102	0.0145	0.0108	0.0134	0.0092	0.0239	0.0082	0.0129	0.0100	0.0136
Schooling	-0.0080	0.0024	0.0141	-0.0059	0.0016	0.0077	-0.0045	-0.0008	0.0012	-0.0056	0.0011	0.0075	-0.0084	-0.0024	0.0011
	0.0108	0.0110	0.0097	0.0051	0.0071	0.0059	0.0108	0.0096	0.0088	0.0051	0.0049	0.0042	0.0053	0.0049	0.0032
R^2		0.6704			0.5816			0.7473			0.5590			0.5907	
Sample Size		1,376			921			161			764			478	
Sample Period		1950-2005			1960-2005			1960-95			1970-2000			1960-85	

Notes: Table reports partial effects at the 25th, 50th, and 75th percentiles along with bootstrapped standard errors. Statistically significant partial effects at the 5% level are emphasized in bold.

Formal statistical tests

We now turn to the nonparametric variable relevance test proposed by Lavergne and Vuong (2000). The above exercise of looking at the statistical significance of the partial effects of each regressor is limited, as insignificance of the partial effects at the quartiles does not suggest overall irrelevance. The formal test of variable relevance specifically tests for the relevance of each regressor (separately) in the model. We report p-values for each variable across each of the education datasets in Table 4. Note that for each dataset, each variable is tested separately from the other variables - each p-value represents a separate test for each variable. In other words, when we test for the relevance of investment, we exclude investment, but keep initial income, population growth, and human capital. Then when we test for the relevance of population growth, we exclude population growth, but keep initial income, investment, and human capital. With the exception of the DD dataset, we find that initial income, physical capital investment, and the population growth variable are statistically relevant variables, with the exception of physical capital investment for CS which has a p-value of 0.1130. We point out that initial income for NSD has a p-value of 0.0840, and is statistically relevant at the 0.1 significance level. Our tests also show that schooling is an irrelevant variable for each of the education datasets.

Our tests also confirm that initial income, investment, and population growth are generally significant and relevant correlates of economic growth. We acknowledge that there is much less ambiguity about the relationship between these variables and economic growth rates, however we highlight that our findings confirm the significance of these variables under arbitrary forms of heterogeneity and model specification. That is, while we find that schooling is not a relevant and significant variable in our model, it is encouraging to provide general and robust results that support the inclusion of these other variables in our growth model.

Before drawing any interpretations or conclusions from the nonparametric test of variable relevance, it is important to address two separate issues. The first is the irrelevance of physical capital investment for CS. Given the sample size of our dataset, we elect not to place much emphasis on a p-value that is just above the 0.1 significance threshold. Second is that the test fails to reject the hypothesis of variable irrelevance for each variable, including the Solow variables, when using the DD sample. Clearly this suggests that the small sample size associated with the DD dataset is leading our test to falsely conclude variable irrelevance (low power). While not particularly surprising, this observation renders interpretation of schooling results for the DD dataset less reliable.

Discussion

Now that we have summarized the results from both the partial effects and formal tests of significance and listed any potential caveats to bear in mind while interpreting our findings, we are in a position to interpret our primary results. Using a nonparametric local-linear estimator and formal test of variable relevance, we find that mean years of schooling is not a statistically relevant variable that is conditionally correlated with higher growth rates. That is, we conclude from our formal nonparametric testing procedure that mean years of schooling is not a robust determinant of economic growth rates. We reiterate that given the general form of our nonparametric regression model, our result is not dependent on any functional form assumptions that might be imposed in traditional parametric growth specifications.

TABLE 4
Significance test p-values for baseline sample

<i>Educational attainment database</i>					
<i>Variable</i>	<i>BL</i>	<i>CS</i>	<i>DD</i>	<i>IIASA-VID</i>	<i>NSD</i>
Initial Income	0.0000	0.0280	0.8010	0.0000	0.0840
Investment	0.0050	0.1130	0.7960	0.0400	0.0190
Population Growth	0.0000	0.0050	0.1540	0.0000	0.0000
Schooling	0.2250	0.6080	0.6510	0.3230	0.3860

Notes: Table reports p-values from the Lavergne and Vuong (2000) tests of the null hypothesis that each of the four variables given in the first column are individually irrelevant variables in our baseline sample. Columns denote which measure of education was included in the model to proxy for human capital. p-values from tests rejecting the null hypothesis at the 5% level are emphasized in bold.

There are several possible explanations for the irrelevance of education. First, there is always the inherent danger of an omitted variable bias. While the nonparametric techniques exploited here are robust to functional form misspecification, they are not robust to omitted variable bias. Reasonable economic theory may suggest the relevance of additional growth determinants not considered here. We note that our nonparametric methods do not allow for the simultaneous inclusion of any and all arbitrary growth determinants one can conjure up since nonparametric estimation requires extraordinarily large sample sizes to achieve the same level of reliability as *correctly specified* parametric models. For a class of *linear* models that do not account for nonlinearities as does the nonparametric framework, yet incorporate a wide range of potential growth determinants, see the Bayesian Model Averaging (BMA) literature (Fernández *et al.*, 2001; Eicher *et al.*, 2007; Durlauf *et al.*, 2008; Eicher *et al.*, 2011).

Second, as was suggested by several authors (Krueger and Lindahl, 2001; Pritchett, 2001) and in part led to the estimation of several new human capital databases in recent years (de la Fuente and Doménech, 2006; Cohen and Soto, 2007; Lutz *et al.*, 2007; Barro and Lee, 2010), there may be potentially large distortions of the estimates due to measurement error. As one might expect, the reliability of data obtained from developing nations may sometimes be low and lack informational content, which may cause severe damage to the regression estimates. Krueger and Lindahl (2001) suggest using samples of countries in which the reliability of the data is greater. It is worthwhile to point out that this is partially addressed via the DD sample and subsequent OECD regression estimates.

Third, years of schooling derived from enrollment rates and census data may in fact provide poor proxies for the stock of human capital within a particular nation. The level of formal education may have little or no significance for the productivity of workers when compared to vocational education, on the job training, or inherent ability. Another possibility is that the level of formal education may affect productivity among certain age cohorts more than others. For example, high-school and college graduates may rely more heavily on their textbook knowledge during their first years in the labor force and depend more on the skills they have acquired on the job after several years. Thus, it may be the case that aggregate human capital measures are insignificant while there still exists a positive impact of schooling on growth.

Fourth, as emphasized by Topel (1999) and Krueger and Lindahl (2001), there is substantial evidence that the time span over which years of schooling is measured matters greatly as there can be little change in a country's average level of schooling in a very short period of time. Thus any estimates of the change in education over shorter time periods (e.g. 5-year intervals) are partly changes associated with measurement errors. When controlling for the change in mean years of schooling over longer time spans, for example 20 years, they find substantial returns to education, suggesting the presence of large externalities to education. We explore this hypothesis in subsequent sections by extending the interval over which our regression model is estimated, and explore the effect of 5-, 10- and 20-year lagged schooling on economic growth.

Fifth, we point out that the results may be highly dependent on the sample of countries included in the regression (Temple, 1998). To explore the effect that different samples may have on our primary results, we consider several alternative segmentations of our initial databases.

V. Robustness checks

In this section, we summarize a battery of robustness checks aimed at dissecting each educational attainment database, controlling for the sample of countries included in the regression, and assessing the reliability of econometric techniques in order to determine whether or not the results presented in the previous section are robust. Specifically, we first re-estimate the nonparametric regression model and perform the variable relevance test using a subsample of observations with outliers removed, use a nonparametric median regression model, and consider an alternative version of the BL dataset that has been corrected for systematic measurement errors (Portela *et al.*, 2010) to fully explore the effect of outliers on our primary results. We further explore the robustness of our primary results by investigating the effect of lagged education on growth rates. We then use all obvious educational attainment subsamples (e.g. OECD/non-OECD samples, male/female samples, etc.) in order to assess whether or not the significance of schooling rests on any particular subgroup. Finally, we consider a consistent sample of observations across each dataset, use a balanced panel, and employ a cross-sectional specification. The results are summarized in Table 5.

We mention that examination of the statistical significance of the partial effects from the local-linear regression models corresponding to each of the robustness checks detailed below yields insignificance (at the quartiles) for the partial effects for the majority of the subsamples. To be clear, the partial effects for each of the subsamples are predominantly insignificant - only in several sporadic cases do we find statistical significance for one of the reported partial effects for a particular education database. Hence, examination of the statistical significance of the partial effects yields a conclusion consistent with that drawn from the Lavergne and Vuong (2000) test of variable relevance for each of the following subsamples or specifications. In what follows, we report only the results from the Lavergne and Vuong (2000) test.¹⁰

Outlier sensitivity

The goal of our first set of robustness checks is to determine whether or not our primary result is robust to the exclusion of outliers. Temple (1998, 2000), for example, emphasizes the potential

¹⁰Unreported regression results are available upon request.

TABLE 5
Significance test p-values for robustness checks

<i>Robustness check</i>	<i>Educational attainment database</i>				
	<i>BL</i>	<i>CS</i>	<i>DD</i>	<i>IIASA-VID</i>	<i>NSD</i>
Outlier sensitivity					
Outlier removal	0.1960	0.6280	0.6310	0.3130	0.3880
Nonparametric LAD	0.0010	0.3193	0.5886	0.0280	0.8218
Portela <i>et al.</i> (2010)	0.6170				
Lagged schooling					
5-year lag	0.5950	0.5810	0.8180	0.5890	0.4790
10-year lag	0.7160	0.5650	0.9410	0.5710	0.4300
20-year lag	0.4380	0.2830	0.7570	0.1390	0.7030
Alternative samples					
OECD nations	0.6950	0.8970	0.6250	0.5500	0.8840
non-OECD nations	0.1990	0.6430		0.5960	0.3560
Males pooled				0.2340	
Males OECD				0.5910	
Males non-OECD				0.5020	
Females pooled	0.2020			0.5760	
Females OECD	0.6650			0.5330	
Females non-OECD	0.1990			0.8070	
Age 25+ pooled	0.5160	0.5430		0.4370	
Age 25+ OECD	0.7330	0.8620		0.5580	
Age 25+ non-OECD	0.5230	0.5790		0.6760	
Age 25+ males pooled				0.2750	
Age 25+ males OECD				0.6300	
Age 25+ males non-OECD				0.5090	
Age 25+ females pooled	0.4260			0.9990	
Age 25+ females OECD	0.7130			0.5400	
Age 25+ females non-OECD	0.4100			0.9990	
Alternative specifications					
10-year intervals	0.3240	0.6540	0.6600	0.4570	0.2130
Small consistent PWT sample	0.9160	0.9080		0.9200	0.9940
Large consistent PWT sample	0.6530	0.7830		0.5070	
African nations	0.3770	0.4850		0.1470	0.1410
non-African nations	0.6320	0.6290		0.2410	0.3710
Balanced panel	0.6310	0.5310	0.3150	0.4230	0.4070
Maasoumi <i>et al.</i> (2007) dataset	0.1850				
Hall and Jones (1999) transformation	0.3540	0.5100	0.7440	0.5080	0.1790
Nonparametric cross-section	0.0960	0.1820	0.4750	0.0250	0.0890

Notes: Table reports p-values from the Lavergne and Vuong (2000) test of the null hypothesis that education is an irrelevant regressor for each robustness check and the relevant education datasets. p-values from tests rejecting the null hypothesis at the 5% level are emphasized in bold.

impact of outliers on growth regression estimates. Temple (2000) stresses the importance of

outlier detection and the use of econometric techniques that are robust to the presence of outliers in order to obtain more precise and reliable estimates. Our nonparametric approach may be sensitive to the presence of outliers since nonparametric methods rely on local weighting that depends on the relative distance between observations. An increase in the relative distance between observations caused by the presence of outliers can potentially influence the estimates obtained from the nonparametric model.

In our first robustness exercise, we use a minimum covariance determinant test and estimator to detect outliers present in our data (Hardin and Rocke, 2004), and proceed by re-running the Lavergne and Vuong (2000) variable relevance test on each human capital dataset with outliers removed.¹¹ The top section of Table 5 shows that the p-values for the test clearly indicate that mean years of schooling is not a relevant variable, even after potential outliers have been removed from the sample.

The second robustness test aimed at assessing the effect outliers may have on our primary result is to run a variable relevance test using a nonparametric least absolute deviation (median) regression model.¹² It is well known that mean regression models are potentially sensitive to outliers, whereas median regression models are less sensitive. Testing for relevance of schooling using the median regression model provides one important method of assessing whether our econometric procedure yields robust conclusions as to the relevance of mean years of schooling in our model. We find that while using the nonparametric least absolute deviation model, we fail to reject the null hypothesis that schooling is an irrelevant regressor for three of the schooling datasets. The test shows that schooling is a relevant variable for both the BL and IASA-VID samples. However, it is important to note here that the partial effects are neither statistically nor economically significant for any of the median regression models.

Our final robustness check assessing the effect of outliers on our primary test results is to use the measurement error corrected version of the BL dataset (Barro and Lee, 2001) constructed by Portela *et al.* (2010). They argue that systematic measurement error in the BL dataset leads to underestimates of mean years of schooling in non-census years - i.e. outliers. They correct for this error using predicted education levels from a multivariate regression model. As can be seen from Table 5, we fail to find relevance of education while using the Portela *et al.* (2010) correction.

We conclude from this section that our conclusions regarding mean years of schooling do not appear to be sensitive to the presence of outliers that may exist in our primary sample.

Lagged schooling

Our second round of robustness checks investigates the potential long run impact of schooling on economic growth. Specifically, we consider using 5-, 10-, and 20-year lags of schooling in place of our schooling variables in the baseline models. Lagging schooling by 5, 10, and 20 years gives an indication of how well past levels of schooling translate into increased growth rates. It is well understood that average years of education change gradually over time, and that a change in the average level of schooling may be slow to manifest itself in improved growth rates.

¹¹See Hartarska *et al.* (2010) for a detailed discussion on the use of the minimum covariance determinant estimator (in a cross-sectional setting) for performing robust nonparametric regression.

¹²We acknowledge that the Lavergne and Vuong (2000) test is theoretically proven for the mean regression model, so we use a test following Zheng (1998) to test for variable relevance in the median regression case.

The second section of Table 5 shows that none of the mean years of schooling measures appear to be relevant variables in any of the lagged human capital specifications. We acknowledge that the p-value for the IASA-VID dataset measured as a 20-year lag is relatively low compared to the other p-values on lagged measures of schooling. This robustness test provides further evidence for our primary result that mean years of schooling is not a relevant variable in our model.

Alternative samples

We now turn to a variety of different education subsets available within the five databases considered here. The third part of Table 5 reports p-values from the Lavergne and Vuong (2000) test for each of the education measurements. We first divide the baseline sample into separate groups based on OECD status to determine whether or not mean years of schooling has a different effect on OECD nations than non-OECD. Focusing solely on OECD nations for all of the datasets also allows us to restrict our sample to a smaller set of countries for which the data are potentially more reliable. It is clear from the reported p-values that none of the available databases show signs of variable relevance for either the OECD or non-OECD samples.

We next consider mean years of schooling measures for males and females separately, for the full sample of countries, OECD countries, and non-OECD countries. Barro (2001), for example, considers differences in the impact of education on growth based on gender, in part because of different degrees of gender discrimination embedded in different countries. Table 5 shows that our primary result that mean years of schooling is irrelevant is robust, and we fail to uncover any statistical relevance of mean years of schooling in our growth model when considering education of males and females separately.

Finally, we reconsider all of our previous subsamples of OECD and non-OECD nations, measured over the total population, males, and females, but for the population aged 25 and above, instead of aged 15 and above. Clearly there may be significant differences in the relation between schooling and growth when schooling is measured over a younger or older working-age population. Regardless of the economic rationale for considering these samples separately, it is clear from Table 5 that mean years of schooling for the population aged 25 and above is not a relevant variable.

Alternative specifications

We finally consider a variety of alternative model specifications that may yield some insight into our finding that mean years of schooling is not a robust and relevant regressor in our nonparametric growth model. We first consider a panel dataset that has been constructed over 10-year intervals instead of 5-year intervals as in our primary sample. The reason for this is two-fold. First, the Cohen and Soto (2007) dataset is originally measured over 10-year intervals, so in the interest of fairness to each education dataset, it is important to consider the Cohen and Soto (2007) (and others) dataset in its original embodiment. Second, measurement error has a potentially weaker impact on the regression estimates as the time interval over which the data are averaged increases, so consideration of a 10-year averaged panel may unearth some signs of variable relevance that may be masked by measurement error in the 5-year panel setting.

Yet, Table 5 clearly shows that we do not identify any statistical relevance when averaging the data over 10-year intervals.

We next construct two datasets that maintain a consistent sample of observations across all of the education databases. We consider this robustness test in order to qualitatively test each of the education databases against each other. The small consistent sample of observations is constructed based on the four largest datasets (i.e. we exclude DD because of insufficient data), which limits our number of pooled country and time observations to 256. The large consistent sample of observations excludes NSD, and is based on BL, CS and IIASA-VID only. The large consistent sample contains 546 pooled country and time observations. As shown in Table 5, maintaining a consistent sample of observations across each of the schooling databases does not alter our primary result that education is irrelevant.

Our next robustness checks are to construct a balanced panel for each dataset, and to consider Sub-Saharan Africa and non Sub-Saharan Africa separately. We investigate whether balancing the panel alters our primary conclusions that mean years of schooling is irrelevant, and as shown in Table 5, our conclusion is reaffirmed. Our rationale for considering African and non-African nations separately stems from recent research suggesting that there are fundamental differences in the growth process when comparing Africa to the rest of the world (Masanjala and Papageorgiou, 2008). Our test of variable relevance fails to uncover statistical relevance of education for either the African or non-African subsamples.

In order to jointly assess whether or not our results are robust to alternative Penn World Table growth rate data and an alternative sample of observations, we estimate the baseline regression model using the data from Maasoumi *et al.* (2007). The growth rate data come from the Penn World Table version 5.6 and the World Bank's World Development Indicators database while the human capital data come from Barro and Lee (2001). Table 5 shows that we fail to find any statistical relevance of education using this alternative dataset.

Our penultimate robustness check is to apply the human capital transformation applied by Hall and Jones (1999) to test whether or not the educational measures show statistical relevance when weighted by a rate of return for each level of educational attainment. We find that regardless of the intuitive appeal of this approach, this data transformation fails to uncover any relevance for the transformed education variable. We conclude that mean years of schooling fails to adequately proxy for human capital even when considering a rate of return for each level of educational attainment.

The final robustness check we consider with our mean years of schooling data is to estimate a nonparametric cross-sectional specification. It is well understood that cross-sectional models can provide relatively robust estimates of the model parameters (Durlauf *et al.*, 2005), since cross-sectional datasets are less influenced by measurement error because they are constructed over larger time horizons. We find that education in the BL and NSD cross-sections is statistically relevant at the 10% significance level, and education in the IIASA-VID cross-section is relevant at the 5% level. Hence, the results found using the Lavergne and Vuong (2000) variable relevance test suggests that, at least for these three datasets, we cannot reject mean years of schooling as a relevant variable in our regression. While this result is encouraging, it is important to analyze the partial effects and their statistical significance in order to determine whether or not the effects of mean years of schooling in the cross-sectional regressions are consistent with prior expectations. We find for the BL dataset, the partial effects are positive and statistically significant taking a mean value of 0.0095, however the magnitude is slightly smaller than we might expect. We did not find any statistical significance for any of the other

cross-sectional datasets. It is important to keep in mind that the sample size (approximately 90 observations, depending on the database) is not large enough to guarantee reliability of the results. Therefore, since we find mixed evidence regarding the relevance of education as a regressor in our nonparametric cross-section model, and little evidence that the partial effects are statistically significant in the nonparametric local-linear regression, we cannot overturn our primary result. However, we acknowledge that the BL data show some signs of promise in the cross-sectional model.

Human capital quality

Having exhausted each of our mean years of schooling databases and having failed to find any strong evidence that mean years of school is relevant, we now turn to an alternative measure of human capital: human capital quality constructed from international test scores. We consider the measure of human capital quality constructed by Hanushek and Kimko (2000), because recent years have seen much discussion focusing on the quality of human capital (Hanushek and Wößmann, 2008), instead of mean years of schooling. The Hanushek and Kimko (2000) data come from international test scores, and are available for a complete cross-section of 75 countries spanning the period 1960-90. Hanushek and Kimko (2000) construct two measures: QL_1 and QL_2 . QL_1 is standardized to the world mean score, and QL_2 is standardized to the mean score from the United States. We consider the log of both measures for two additional cross-sectional robustness checks.

The p-values for QL_1 and QL_2 from the Lavergne and Vuong (2000) relevance test are 0.0640 and 0.0750, respectively. We find for QL_1 that the local-linear partial effects are positive and statistically significant at the 5% level at the mean, 25th, 50th, and 75th percentiles. While this result seems to mirror the BL cross-sectional result, we point out that the partial effects are of much larger magnitude when using the Hanushek and Kimko (2000) data. The partial effects are 0.0204, 0.0320, and 0.0380 at the 25th, 50th, and 75th percentiles. All partial effects for the other regressors have magnitude, sign, and significance consistent with a priori expectations. We obtain similar partial effect results for the QL_2 measure. Therefore, the human capital quality data appear to perform admirably in our nonparametric model.

We caution, however, that these results are based on a nonparametric cross-sectional model, and 75 observations may not be large enough to obtain fully reliable results. We emphasize that the model provides results completely consistent with prior expectations, and appears to outperform each of the mean years of schooling datasets. At the very least, this exercise seems to suggest that further research into quality measures or cognitive ability may prove fruitful.

VI. Conclusion

While there is undeniably ample intuition and theoretical support for the importance of human capital in growth specifications, we fail to find robust empirical support for the inclusion of human capital as proxied by mean years of schooling in our baseline growth regression specification. While the previous empirical growth literature has failed to reach a consensus as to the significance and relevance of educational attainment in growth models, we point out that much of the discrepancy can be attributed to differences in model specification and human capital measurement. Using five leading educational attainment databases in conjunction with

nonparametric econometric techniques that are robust to functional form misspecification, and employing a barrage of robustness checks addressing concerns over both data structure and measurement, we fail to find adequate empirical support for the statistical significance of mean years of schooling. Rather, we find that, in general, estimates of the impact of mean years of schooling bear little statistical significance in terms of the estimated partial effects and a formal nonparametric test of variable relevance (Lavergne and Vuong, 2000).

Our results have several implications. First, we find robust evidence that previous empirical results identifying positive economic effects of educational attainment are most likely due to functional form misspecification. Second, we provide evidence that outliers and differences in the sample of observations are not the cause of our primary finding that educational attainment is a statistically irrelevant variable. Third, our finding that educational achievement shows relevance and statistical significance while educational attainment does not suggests that educational attainment may not adequately represent human capital in a growth regression.

Our results naturally lead us to the following question. What should be done in order to adequately represent and model human capital in growth regressions? There are several possible answers to this question. First, it is plausible that the models considered here failed to include variables measuring all significant determinants of economic growth. Re-examination of additional potential growth determinants, perhaps through a nonparametric lens, may reveal additional variables that significantly influence the growth process and therefore resolve any omitted variable problem. Second, further research into alternative measures of human capital may be warranted. Our nonparametric models show that the practical consequences of using different education databases are modest, contrary to the findings of previous research (de la Fuente and Doménech, 2002; Cohen and Soto, 2007; Lutz *et al.*, 2007). This suggests human capital may not be adequately represented by mean years of schooling, thereby providing support for theories speculating as to the quality of years of schooling obtained in any given country relative to other countries, and whether or not vocational training has a greater effect on economic growth than formal years of schooling.

Additional research has focused on the quality of human capital (Wößmann, 2002) and the effect of cognitive skills on economic growth (Hanushek and Kimko, 2000; Hanushek and Wößmann, 2008). Our nonparametric cross-sectional exploration of the Hanushek and Kimko (2000) data shows a significant and robust effect of human capital quality on economic growth rates. Recently developed techniques, such as sophisticated projection techniques developed through the construction of existing education databases, might be applied to measure changes in mean education achievement, to help develop these data into larger and more widely available datasets.

References

- [1] Barro, R. J. (1991). 'Economic growth in a cross-section of countries', *Quarterly Journal of Economics*, Vol. 106, pp. 407-443.
- [2] Barro, R. J. (2001). 'Human capital and growth', *American Economic Review: Papers and Proceedings*, Vol. 91, pp. 12-17.
- [3] Barro, R. J. and Lee, J. (1993). 'International comparisons of educational attainment', *Journal of Monetary Economics*, Vol. 32, pp. 363-394.

- [4] Barro, R. J. and Lee, J. (1996). 'International measures of schooling years and schooling quality', *American Economic Review: Papers and Proceedings*, Vol. 86, pp. 218-223.
- [5] Barro, R. J. and Lee, J. (2001). 'International data on educational attainment: updates and implications', *Oxford Economic Papers*, Vol. 3, pp. 541-563.
- [6] Barro, R. J. and Lee, J. (2010). 'A new data set of educational attainment in the world, 1950-2010', NBER Working Paper 15902.
- [7] Bils, M. and Klenow, P. J. (2000). 'Does schooling cause growth?', *American Economic Review*, Vol. 90, pp. 1160-1183.
- [8] Cameron, C. A. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*, Cambridge University Press, New York, NY.
- [9] Cohen, D. and Soto, M. (2007). 'Growth and human capital: good data, good results', *Journal of Economic Growth*, Vol. 12, pp. 51-76.
- [10] de la Fuente, A. and Doménech, R. (2002). 'Educational attainment in the OECD, 1960-1995', CEPR DP 3390.
- [11] de la Fuente, A. and Doménech, R. (2006). 'Human capital in growth regressions: how much difference does data quality make?', *Journal of the European Economic Association*, Vol. 4, pp. 1-36.
- [12] Duffy, J. and Papageorgiou, C. (2000). 'The specification of the aggregate production function: a cross-country empirical investigation', *Journal of Economic Growth*, Vol. 5, pp. 83-116.
- [13] Durlauf, S. N. and Johnson, P. A. (1995). 'Multiple regimes and cross-country growth behavior', *Journal of Applied Econometrics*, Vol. 10, pp. 365-384.
- [14] Durlauf, S. N., Johnson, P. A. and Temple, J. R. W. (2005). 'Growth econometrics', *Handbook of Economic Growth*, Vol. 1, pp. 555-677.
- [15] Durlauf, S. N., Kourtellos, A. and Minkin, A. (2001). 'The local Solow growth model', *European Economic Review*, Vol. 45, pp. 928-960.
- [16] Durlauf, S. N., Kourtellos, A. and Tan, C. M. (2008). 'Are any growth theories robust?', *Economic Journal*, Vol. 118, pp. 329-346.
- [17] Eicher, T. S., Papageorgiou, C. and Raftery, A. E. (2011). 'Default priors and predictive performance in Bayesian model averaging, with application to growth determinants', *Journal of Applied Econometrics*, Vol. 26, pp. 30-55.
- [18] Eicher, T. S., Papageorgiou, C. and Roehn, O. (2007). 'Unraveling the fortunes of the fortunate: an iterative Bayesian model averaging (IBMA) approach', *Journal of Macroeconomics*, Vol. 29, pp. 494-514.
- [19] Fernández, C., Ley, E. and Steel, M. F. J. (2001). 'Model uncertainty in cross-country growth regressions', *Journal of Applied Econometrics*, Vol. 16, pp. 563-576.

- [20] Gemmell, N. (1996). 'Evaluating the impacts of human capital stocks and accumulation on economic growth: some new evidence', *Oxford Bulletin of Economics and Statistics*, Vol. 58, pp. 9-28.
- [21] Hall, R. E. and Jones, C. I. (1999). 'Why do some countries produce so much more output per workers than others?', *Quarterly Journal of Economics*, Vol. 114, pp. 83-116.
- [22] Hanushek, E. A. and Kimko, D. D. (2000). 'Schooling, labor-force quality, and the growth of nations', *American Economic Review*, Vol. 90, pp. 1184-1208.
- [23] Hanushek, E. A. and Wößmann, L. (2008). 'The role of cognitive skills in economic development', *Journal of Economic Literature*, Vol. 46, pp. 607-668.
- [24] Hardin, J. and Roche, D. M. (2004). 'Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator', *Computational Statistics and Data Analysis*, Vol. 44, pp. 625-638.
- [25] Hartarska, V., Parmeter, C. F., Nadolynak, D. and Zhu, B. (2010). 'Economies of scope for microfinance: differences across output measures', *Pacific Economic Review*, Vol. 15, pp. 464-481.
- [26] Henderson, D. J. (2010). 'A test for multimodality of regression derivatives with an application to nonparametric growth regressions', *Journal of Applied Econometrics*, Vol. 25, pp. 458-480.
- [27] Henderson, D. J., Papageorgiou, C. and Parmeter, C. F. (2012). 'Growth empirics without parameters', *Economic Journal*, Vol. 122, pp. 125-154.
- [28] Heston, A., Summers, R. and Aten, B. (2011). 'Penn World Table Version 7.0', Technical report, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- [29] Islam, N. (1995). 'Growth empirics: a panel data approach', *Quarterly Journal of Economics*, Vol. 110, pp. 1127-1170.
- [30] Kalaitzidakis, P., Mamuneas, T. P., Savvides, A. and Stengos, T. (2001). 'Measures of human capital and nonlinearities in economic growth', *Journal of Economic Growth*, Vol. 6, pp. 229-254.
- [31] Krueger, A. and Lindahl, M. (2001). 'Education for growth? Why and for whom?', *Journal of Economic Literature*, Vol. 39, pp. 1101-1136.
- [32] Lavergne, P. and Vuong, Q. (2000). 'Nonparametric significance testing', *Econometric Theory*, Vol. 16, pp. 576-601.
- [33] Liu, Z. and Stengos, T. (1999). 'Non-linearities in cross country growth regressions: a semiparametric approach', *Journal of Applied Econometrics*, Vol. 14, pp. 527-538.
- [34] Lutz, W., Goujon, A., Samir, K. C. and Sanderson, W. (2007). 'Reconstruction of populations by age, sex, and level of educational attainment for 120 countries for 1970-2000', *Vienna Yearbook of Population Research*, pp. 193-235.

- [35] Maasoumi, E., Racine, J. S. and Stengos, T. (2007). 'Growth and convergence: a profile of distribution dynamics and mobility', *Journal of Econometrics*, Vol. 136, pp. 483-508.
- [36] Mankiw, N. G., Romer, D. and Weil, D. N. (1992). 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics*, Vol. 107, pp. 407-437.
- [37] Masanjala, W. H. and Papageorgiou, C. (2008). 'Rough and lonely road to prosperity: a reexamination of the sources of growth in Africa using Bayesian model averaging', *Journal of Applied Econometrics*, Vol. 23, pp. 671-682.
- [38] Mincer, J. (1974). *Schooling, Experience, and Earnings*, Columbia University Press, New York, NY.
- [39] Minier, J. (2007). 'Nonlinearities and robustness in growth regressions', *American Economic Review: Papers and Proceedings*, Vol. 97, pp. 388-392.
- [40] Mitchell, B. R. (1982). *International Historical Statistics, Africa and Asia*, Second Edition, New York University Press, New York, NY.
- [41] Nehru, V., Swanson, E. and Dubey, A. (1995). 'A new database on human capital stock in developing and industrial countries: sources, methodology and results', *Journal of Development Economics*, Vol. 46, pp. 379-401.
- [42] Nonneman, W. and Vanhoudt, P. (1996). 'A further augmentation of the Solow model and the empirics of economic growth for OECD countries', *Quarterly Journal of Economics*, Vol. 11, pp. 943-953.
- [43] Portela, M., Alessie, R. and Teulings, C. (2010). 'Measurement error in education and growth regressions', *Scandinavian Journal of Economics*, Vol. 112, pp. 618-639.
- [44] Pritchett, L. (2001). 'Where has all the education gone?', *World Bank Economic Review*, Vol. 15, pp. 367-391.
- [45] Pritchett, L. (2006). 'Does learning to add up add up? The returns to schooling in aggregate data', *Handbook of the Economics of Education*, Vol. 1, pp. 635-695.
- [46] Racine, J. S. and Li, Q. (2004). 'Nonparametric estimation of regression functions with both categorical and continuous data', *Journal of Econometrics*, Vol. 119, pp. 99-130.
- [47] Sala-i-Martin, X., Doppelhofer, G. and Miller, R. I. (2004). 'Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach', *American Economic Review*, Vol. 94, pp. 813-835.
- [48] Savvides, A. and Stengos, T. (2009). *Human Capital and Economic Growth*, Stanford University Press, Stanford, CA.
- [49] Schoellman, T. (2012). 'Education quality and development accounting', *Review of Economic Studies*, Vol. 79, pp. 388-417.
- [50] Sianesi, B. and Van Reenen, J. (2003). 'The returns to education: macroeconomics', *Journal of Economic Surveys*, Vol. 17, pp. 157-200.

- [51] Temple, J. R.W. (1998). 'Robustness tests of the augmented Solow model', *Journal of Applied Econometrics*, Vol. 13, pp. 361-375.
- [52] Temple, J. R. W. (2000). 'Growth regressions and what the textbooks don't tell you', *Bulletin of Economic Research*, Vol. 52, pp. 181-205.
- [53] Temple, J. R. W. (2001). 'Generalizations that aren't? Evidence on education and growth', *European Economic Review*, Vol. 45, pp. 905-918.
- [54] Topel, R. (1999). 'Labor markets and economic growth', *Handbook of Labor Economics*, Vol. 3C, pp. 2943-2984.
- [55] Wößmann, L. (2002). *Schooling and the Quality of Human Capital*, Springer, Berlin.
- [56] Zheng, J. X. (1998). 'A consistent nonparametric test of parametric regression models under conditional quantile restrictions', *Econometric Theory*, Vol. 14, pp. 123-138.