THE IMPORTANCE OF EDUCATIONAL QUALITY IN DETERMINING STATE INCOME GROWTH, INCOME INEQUALITY AND POVERTY

by

Jean Le Moullec

Submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

January 2011

Abstract[*]

In the past decade, studies have begun investigating the impacts of educational quality, as measured by standardised test scores, on income per capita growth. Past studies of growth have ignored quality or assumed that its effects on growth are negligible. Recent examinations that have included controls for both the quantity and the quality of schooling have found that quality of education may not only be important, but that it may even have a larger impact on growth than the quantity of education. Using a panel dataset of the U.S. states, this study continues the investigation of the impacts of educational quality on growth and also explores if quality has effects on income inequality and the poverty rate. While controlling for the quantity of education, the quality of education (as measured by test scores) is found to have a statistically significant positive impact on income growth. However, due to issues of unobservable state specific fixed effects and serial correlation, the result is sensitive to the estimation method used. For income inequality and poverty, once state fixed effects are accounted for, quality has no impact on either of these two macro level characteristics.

[*] I would like to thank my supervisor Marco Cozzi for his helpful guidance and support throughout the production of this essay.

Table of Contents

List of Tables

Introduction

For decades, the study of the effects of education has been at the centre of many great debates. These studies range from trying to calculate the monetary returns of an additional year of education to estimating the impact of increasing education on crime rates. Many questions posed about education are of great importance to people, communities, governments, countries, continents, and the world as a whole. Because education is so crucial to so many issues, the findings concerning the effects of education on wages, health, well-being, and so forth have vast repercussions on the various educational policies adopted by governments.

Perhaps the issue that has received the most attention is the one of education, or more precisely human capital, and economic growth. Any policy that can result in even a small increase in the growth rate can have drastic changes on the standard of living in the long run. Incorporating human capital in models of economic growth was very popular throughout the 1990s. Two interesting issues of these earlier studies are that (i) human capital, often proxied by some form of years of schooling, does not always prove to be a significant determinant of growth depending on how the sample of countries is chosen, specifically if we restrict our view to developed countries (Coulombe and Tremblay 2006), and (ii) the quality of human capital is rarely addressed.

With regard to the first issue, there is more than one possible explanation for why this occurs. One possibility is the fact that countries have different levels of educational quality, which in turn corresponds to the second issue. Years of schooling simply gives a notion of the quantity of human capital and ignores issues of quality. For example, even though Kazakhstan has a similar level of average years of schooling as France[2], many would likely expect that a

[2] According to the Barro schooling data from the World Bank, in 2005 Kazakhstan's average years of schooling for the population 15+ years was 10.1 and France's was 9.9.

secondary diploma from the latter is superior to that of the former. This second issue can be resolved if we believe that any differences in quality of education are minute and are outweighed by differences in the quantity of education. This latter idea is very reasonable when one thinks of countries such as the U.S. and the U.K., but has less credibility when one compares countries such as Kazakhstan and France. Even within a country, one can imagine comparing the quality of education at Harvard University versus the quality at a generic community college. Few people would ever believe they are the same. Furthermore, given the plethora of variables that are included as controls in growth models, it seems unreasonable not to address issues concerning the quality of education or, in other words, the quality of human capital.

Recently, researchers have begun examining the effect of quality and not only quantity of education where quality is typically controlled for using performance on a set of standardized tests. As of yet, however, these analyses have restricted their work to cross-country investigations and have primarily only looked at the effects of quality on economic growth. In addition, panel-data investigation on the issue of quality, which allows for more sophisticated econometric methods, has been limited. This may be due to the lack of richness in cross-country standardized testing data.

In this essay, I will examine the importance of the quality of education on not only economic growth, but on income inequality and poverty rates as well. Moreover, instead of a cross-country cross-sectional study, this analysis will focus on the fifty U.S. states using panel-data estimation procedures for the 1990 to 2007 period.

The desire to observe and control for the quality of human capital in growth models is similar to the need to account for individual ability in wage determination models. Though some may believe the former is needless, few would argue the necessity of the latter, yet one is more or less the macroeconomic equivalent of the other. Not surprisingly, like ability in a wage model, one way to control for quality of education is to incorporate standardized test scores. This is exactly what is done by Hanushek and Kim (1995). Hanushek and Kim make use of several international tests that have been administered in many countries since the 1960s[3]. These test scores allow the researchers to factor in the quality of education and not only the quantity. Using test scores to account for ability in wage determination models is not a new concept. However, aggregating scores to attain a score for a defined region and using this score at the macroeconomic level is. Since Hanushek and Kim (1995), others have re-examined the growth question while making use of this method to control for the quality of human capital (see Hanushek and Kimko (2000); Barrow (2001); and Jamison, Jamison, and Hanushek (2007)).

Although this method for controlling for the quality of education by standardized test scores has been primarily used in growth models, its application by no means ends there. Considering how education is front and centre in so many different issues, it seems natural to start trying to incorporate quality, as measured by standardized test scores, into the investigation of numerous other macroeconomic phenomena, such as poverty and income inequality which are studied in this essay.

One problem with using the international test scores to examine the effects of the quality of human capital is that the tests are, up until recently, administered infrequently and the countries participating are not always numerous or the same. To address this, Hanushek and Kim create a composite of all the test scores for different international tests in the subjects of mathematics and science. They develop a method that standardizes all test scores throughout the 1960-1990 period. The quality score they create then allows for a cross-country cross-sectional analysis of the quality of education.

[3] See Hanuskek and Kim (1995) for a presentation and discussion of international testing data.

The drawback to the cross-country analysis with the quality score is directly related to how the score was created. To have a large enough sample, one must include countries that may have only participated in a few tests (in different testing years) over a twenty or thirty year period. For this reason, applying panel-data analysis to investigate the issue dynamically becomes very difficult. The ideal data would have countries participating in every testing year. This data does not yet exist at the international level. However, it does exist at the state level in the United States. This state level panel-data is precisely what is used in this essay.

The National Association for Educational Progress (NAEP), which is part of the U.S. Department of Education, has been conducting standardized tests across the fifty states in multiple subject areas on a regular basis since 1990. The advantage with this data is that, because the vast majority of states participated in the standardized testing every testing year, little has to be done to the test scores to allow for a cross-state panel-data analysis of the effects of the quality of education on economic growth, poverty, and income inequality, where quality is proxied by the test scores. One advantage that the cross-country analysis has over the cross-state analysis is that there exists much more heterogeneity between countries of varying levels of development than between states. Nonetheless, there are still some economic and social differences across the U.S. states to permit us to hopefully gain more knowledge on the impacts of the quality of education in several macroeconomic issues.

The remainder of this essay will begin with a survey of the literature on the quality of education, followed by a presentation of the data used, a discussion of the various estimation results, and then some concluding remarks.

Literature Survey

The idea to control for quality of education and not just quantity in growth models was started by Hanushek and Kim (1995). In their essay, as mentioned in my introduction, they construct a measure for the quality of education using international standardized test scores and include it in a typical growth regression[4]. Their results show that a one standard deviation increase in their quality measure increases growth by a whole percentage point, which is larger than the effect of increasing average years of schooling by eight years. Interestingly though, when test scores are regressed on measures of schooling resources, little variation in test scores can be explained. This finding, coupled with the finding of the importance of quality, results in a policy dilemma for governments. If slightly improving the quality of education (in this case test scores) can greatly increase growth, what are governments to do if the policies readily available to them (i.e. increase education expenditures) have essentially no impact on test scores? Though this conundrum is obviously of immense importance it is not investigated in this essay.

Since 1995, Hanushek, with others, has written several papers on growth using a quality measure (see Hanushek and Kimko (2000), Jamison, Jamison, and Hanushek (2007), Hanushek and Wobmann (2007)). Though the idea dates back to the paper by Hanushek and Kim from 1995 (which was not published), the popularity of this approach grew faster after Hanushek and Kimko's essay *Schooling, Labor-Force Quality, and the Growth of Nations* (AER 2000). In this research, Hanushek and Kimko perform an analysis similar to the one done in the 1995 essay and find that the magnitude of the effects of quality on growth is, again, that a one standard deviation increase in quality leads to a one percentage point increase in growth with this increase

[4] Hanushek and Kim actually investigate two specifications of quality using the test scores. One adjusts all scores to have an average of fifty and the other is constructed using a more sophisticated process where scores are adjusted based on the U.S. international testing scores adapted for the U.S. national time pattern of scores on the NAEP tests (see Hanushek 2010 *The high cost of low educational performance* Annex A for a detailed explanation of how this second measure is constructed).

approximately the same as a nine year increase in average years of schooling. However, the authors point out that the effects of quality are somewhat unbelievably large and should thus be interpreted with caution.

Table 1 shows new cross-country estimates for the effects of educational quality (i.e. human capital quality) on growth for the 1995-2007 period. I find that, although quality is important, its impact has severely diminished in this more recent data that includes a larger set of countries than in Hanushek and Kim (1995).

Table 1: Quantity vs. Quality: Cross-country OLS regression - Cross-sectional data

Explaining average annual growth in GDP per capita (2000 US$, percentage)	[1]	One Std. Impact	[2]	One Std. Impact
population growth rate (percentage)	-0.48***	-0.48	-0.40**	-0.40
	[0.17]		[0.18]	
GDP per capita in 1995 (1000s US$, 2000=100)	-0.09***	-0.90	-0.10***	-1.00
	[0.02]		[0.02]	
years of schooling	0.24**	0.48	0.19*	0.38
	[0.10]		[0.10]	
quality of education[1]			0.04**	0.26
			[0.02]	
Constant	2.05**		0.27	
	[0.80]		[1.21]	
Observations	62		62	
R-squared	0.47		0.49	

* p<0.10, ** p<0.05, *** p<0.01

1 - Quality of education is the international test scores scaled to have an average of fifty (as in Hanushek and Kim 1995).
Note: Robust standard errors are reported in square brackets.

The reason I chose 1995 as the starting year is because 1995 is the year when the Third International Mathematics and Science Study (TIMSS) began testing. Unlike other datasets, the testing administered by TIMSS is prepared in a way such that the resulting data is comparable across countries. To do this, much work is put into ensuring that testing is standardized across languages, countries, and cultures (TIMSS Quality Assurance in Data Collection 1996). TIMSS conducted testing in 1995, 1999, 2003, and 2007. All data used for this regression has been attained from the World Bank website. To be included in the regression, a country needed only to have a score for one of the testing years. To remain consistent with the cross-state analysis presented later, the quality of education is proxied by grade eight test scores in mathematics. The dependent and independent variables are all period averages. More information concerning this data and the participating countries can be found in the appendix of this essay (see tables A1-A4).

Table 1 shows that all explanatory variables have significant coefficients and the signs on the coefficients display the expected relationships commonly found in growth regressions: higher population growth translates to lower growth; higher initial GDP per capita leads to lower growth (implying the conditional convergence hypothesis holds); more human capital (as proxied by years of schooling) leads to higher growth; and higher quality of human capital (i.e. quality of education) also leads to higher growth. With respect to human capital, when quality is included in the model the magnitude of the coefficient on years of schooling diminishes, as found in past studies. A one standard deviation increase in years of schooling, when quality of education is excluded, is associated with a 0.48 percentage point increase in average annual growth in income per capita. When we control for quality, the impact associated with years of schooling falls to a 0.38 percentage point increase (for a one standard deviation increase). This is slightly larger to what is found in Hanushek and Kimko (2000). They find that, after controlling for quality, a one standard deviation increase in years of schooling raises growth by 0.26 percentage points. For quality of education, here a one standard deviation increase corresponds to a 0.26 percentage point increase in average yearly growth, which is much smaller than the somewhat unbelievably high baseline estimate of a 1.46[5] point increase found by Hanushek and Kimko (2000). As previously mentioned, the authors themselves point out that the size of the impact they attain should be interpreted with caution. Even if the true effects are 0.26[6], the increase is still not a trivial one. Compounded over many years, small increases in the growth rate will result in significant changes in the level of income per capita. Magnitude aside, what should be taken away from this little exercise is that the quality of education still appears to be an important determinant of growth at the cross-country level.

Hanushek and Kimko (2000) also explore the question of causality between test scores and economic growth. It could just as well be that high growth rates are causing higher test scores, or more believably, high growth rates and high test scores are correlated to some omitted variable which is causing both. To shed light on this, the researchers estimate a traditional Mincerian wage determination model for a sample of male immigrants in the U.S. where they can identify whether an immigrant was educated in the U.S., in their home country, or in both. If origin country test scores are just the outcome of some country specific factor and have no bearing on labour force quality, then there should be no returns to an immigrant's wage of coming from a country with high test scores (or in other words, coming from a country with better schools and hence more productive workers). If there are statistically significant returns to wages, then the quality of education variable would appear to be able to account for differences in individual productivity and thus, at the macroeconomic level, be able to indicate more productive human capital across countries. If this holds, then there is reason to believe that quality of education is causing higher growth rates and not the other way around. Furthermore, when the sample is restricted to immigrants educated in the U.S., if country quality of education

---

[5] This number is taken from the baseline model estimates of the Hanushek and Kimko paper. The one percentage point increase mentioned earlier comes from models they estimate later in their paper.
[6] For sake of argument I am implicitly assuming causation.

(i.e. test scores) is implying productivity, then we should observe no significant effects for schooling quality since all men in this sample have accumulated their human capital in the same country[7].

For those entirely educated in their home country, Hanushek and Kimko's regression results show that a one unit increase in quality of education (i.e. higher test scores) increases wages by 0.21 percent. This result gives reason to believe that test scores are, one, able to indicate the quality of human capital in a worker and, two, causing growth. The results for the sample of immigrants entirely educated in the U.S. yields insignificant coefficients giving more evidence that test scores are indicating productivity and not just home country specific effects (Hanushek and Kimko 2000).

In Barro (2001), Barro performs a growth analysis using test scores to account for differences in human capital quality. He regresses growth on combinations of science score, mathematics score, reading score, total score, and proportion of males with secondary education or higher. Barro's results show that the quality of human capital, proxied by these test scores, has, again, a much larger impact on growth than the measure for the quantity of human capital.

One significant drawback of these initial studies on the quantity versus the quality of education is that, because of data limitations, the analyses are restricted to cross-country cross-sections and thus cannot make use of any panel estimation methods that allow for more in depth investigations. To get around this problem, Coulombe and Tremblay (2006) use the results of the 1994-1998 Adult Literacy Survey, which is used to evaluate literacy proficiency of those between 16 and 65 years of age, for 14 OECD countries. With this data, they construct a synthetic time-series for the period of 1960 to 1995 using the age distribution of the test scores. With the data on age distributions, the average literacy scores for those 17-25 for the initial year

[7] This implicitly assumes that quality of education is the same across regions in the United States.

of each five year period (1960, 1965, etc) are used as proxies for the productivity of human capital within a country and are in turn used to measure the initial relative investment in human capital. They then perform panel estimation with the data which is reduced to seven five-year periods. As Coulombe and Tremblay point out, two problems with their variable for the quality of human capital are that it does not address migration flows and assumes that an individual's level of human capital is constant over time. If period migration flows are large then a country's average test score will not accurately capture cross-country differences in the quality of human capital since an immigrant's level of human capital will be reflected by his home country score and not the score of the country being observed. This problem of migration flows is also an issue in the state data used in this essay. As for their second problem, the assumption that an individual's level of human capital is constant can result in an overestimation of the period investments in human capital because the test data is taken from the end period of the analysis. Thus, if literacy ability increases over one's lifetime through on-the-job learning, or through any other facet, then it is quite possible that older individuals will systematically score higher than younger individuals on the 1994-1998 tests and will then appear to have come from a period that had higher investment in human capital.

Using a panel estimation procedure similar to that used by Barro and Sala-i-Martin (2004) and various other techniques, Coulombe and Tremblay find that their human capital quality (as proxied by the prose score, document score, quantitative score, or overall literacy score[8]) has a positive and significant impact on the growth rate of GDP per capita[9]. Interestingly, the magnitude and significance of the quality measure remained almost constant regardless of which of the four scores was used as the proxy. The authors believe that this means

[8] The test examined three domains: prose, document, and quantitative, and participants were given a score in each domain. The literacy score was taken as the average of the three scores.
[9] Their model also controlled for country fixed effects and time dummies.

that the information concerning human capital quality that is imbedded in the scores is quite comparable across domains. The authors also run their model using years schooling data from Barro and Lee (2001) and de la Fuente and Doménech (2002) as the human capital indicator and find that neither datasets resulted in significant coefficients for human capital. This finding leads them to conclude that the literacy scores contain more information about human capital than years of schooling, yet they point out that the two measures are not directly comparable since the first is a quality measure (or as they say, signifies investment in human capital) and the other is a quantity (or a stock). Unfortunately, and somewhat puzzling, Coulombe and Tremblay do not follow Hanushek and Kimko (2000) and include both measures simultaneously.

The estimation method used in this analysis resembles most the one adopted by Coulombe and Tremblay (2006) except that I include both measures of human capital, quality and quantity, in the regression model.

Another paper that uses cross-country panel-data (for the 1960-2000 period) but includes both quantity and quality measures for human capital is Jamison, Jamison, and Hanushek (2007). However, this study focuses on examining the process by which the quality of human capital affects growth rates and if increases in human capital quality can decrease the rate of infant mortality. The three processes considered are (i) quality expands the level of the production function by changing country fixed effects, (ii) quality raises the marginal effect of an additional year of schooling, and (iii) quality raises the rate of technological progress. Their analysis begins by estimating a similar model as in Hanushek and Kimko (2000). In addition, they estimate a model where controls for openness, fertility rate, and land area in the tropics are included. The results show that a one standard deviation increase in the quality measure (as measured by mathematics test scores[10]) increases the growth rate by 0.87 points for the first model and 0.45 points for the model with more controls. The authors point out that the first estimate is quite comparable to the full 1.0 percentage point increase previously found but that the latter estimate, at 0.45, is much more believable. Their panel estimation results show that a one standard deviation increase in the test score variable increases growth in income per capita by 0.5-0.9 percent, depending on the model used.

As for which of the three mechanisms of how the quality of education affects growth, the results of their analysis lead Jamison, Jamison, and Hanushek to conclude that quality affects growth in per capita income through the rate of technological progress[11]. The results concerning the infant mortality rate show that a one standard deviation in the quality variable is associated with a 0.6 percent drop in the infant mortality rate. This result is important to us because the finding that quality has a significant effect on the infant mortality rate opens the door to other avenues of research on the impacts of the quality of education, such as on income inequality and poverty.

The Data

The data used in this analysis comes primarily from three places: the U.S. Bureau of Labor Statistics (BLS), the U.S. Bureau of Economic Analysis (BEA), and the U.S. Current Population Survey (CPS). The standardized tests scores are taken from the National Assessment of Educational Progress (NAEP). To proxy for the quality of education at the state level, I make use of the grade 8 state average scores in mathematics. Test scores for other subjects are available. Standardized testing is done on several subject areas for three different grade levels;

[10] The quality measure in this essay is constructed in a similar fashion as in Hanushek and Kimko (2000) yet only uses mathematics scores instead of scores for both mathematics and science.
[11] The authors adopted a "meta production function" approach and multi level modeling techniques, a maximum likelihood procedure, to estimate their model.

grade 4, grade 8, and grade 12, or students aged 9, 13, and 17 years respectively. The subject areas tested are Civics, Economics, Geography, Mathematics, Music, Reading, Science, U.S. History, Visual Arts, and Writing. Testing is performed uniformly using the same test booklets making test scores perfectly comparable across states (NAEP Overview 2010). This is a great advantage over many other sources of testing data since testing is done from a national perspective, and so comparison at lower levels of aggregation (in this case state level) is straightforward.

The ideal measure for educational quality would make use of test scores across all subjects. The reason only mathematics test scores for grade 8 students are used is because it was the subject and grade that had data for the longest available period. Only using mathematics scores, as is done in Jamison, Jamison, and Hanushek (2007), poses no real problem since test scores across subjects and grades are highly correlated (see Table 2).

Table 2: Correlation table for standardized test scores by subject

	math 4	math 8	reading 4	reading 8	science 4	science 8
math 4	1	0.95	0.90	0.85	0.80	0.81
reading 4	0.90	0.90	1	0.95	0.92	0.89
reading 8	0.85	0.90	0.95	1	0.92	0.95
science 4	0.80	0.84	0.92	0.92	1	0.94
science 8	0.81	0.89	0.89	0.95	0.94	1

Note: The above correlation table is for the period of 2002 to 2005 and excludes Alaska, Iowa, Kansas, Nebraska, New York, and Pennsylvania.

Table 2 shows the correlation for the 2002 to 2005 period for 44 of the 50 states, as scores for Alaska, Iowa, Kansas, Nebraska, New York, and Pennsylvania were not available for at least one of the subject and grade scores. The 2002 to 2005 period was the only period where testing had occurred at least once for all three subjects and all three grades. As can be seen in the table, the scores are highly correlated with correlations ranging from 0.80 to 0.95. Because of this high correlation, using any of the scores as a proxy for educational (or human capital) quality would suffice. Therefore, using the subject and grade with the most data points, in this case mathematics and grade 8, is the best option.

From the BLS, I use the annual state unemployment rate. From the BEA, I acquire the state population growth rate and the growth in per capita income. With the CPS, I derive state level variables using the state indicators attached to every respondent. Surprisingly, state level data on population characteristics is very hard to come by. This is why micro level data, instead of the preferred macro level data, is used to attain state level variables[12]. From the CPS, I attain variables for race (specifically black or white), years of schooling and educational achievement, poverty rates, and personal income (from which I calculate a gini coefficient). The CPI data used for deflation (base 1999) is also from the CPS[13]. More detail of the variables used and how they are constructed is provided in the appendix (see Table A5). Table 3 further describes the variables mentioned above.

Table 3: Variables, descriptions, and sources

Variable	Description	Source
inc gr	Growth rate in personal income per capita (percentage)	BEA
inc1990	Income per capita at the start of the period (thousands)	BEA
pop gr	Population growth rate	BEA
ump rate	Unemployment rate	BLS
CPI (1999=100)	National level consumer price index with base in 1999	CPS
gini	Gini coefficient for personal income	CPS
white	Percent of population who identifies their race as white	CPS
black	Percent of population who identifies their race as black	CPS
bach	Percent of population with a Bachelor's degree or more	CPS
poverty	Percentage of the population living in poverty	CPS
school	Imputed years of schooling	CPS
qed	Quality of Education (QED): Percentage point difference from state score and cross-state mean, in mathematics for a given testing year	NAEP

[12] Because this dataset is not specifically designed at the macro level, there exists the possibility of sample biases. However, since the dataset is quite large, approximately 3.5 million observations for the 1990-2007 period, it is assumed that any potential sampling bias is mitigated making the sample values valid estimates for the true population values.

[13] State specific CPIs would have been preferred yet none were available that covered the entire period of study.

With respect to the quality of education variable, as measured by mathematics scores, the reason I take the percentage point difference from the state score and the testing year cross-state mean is twofold. Firstly, taking the difference from the mean adjusts for test difficulty. If the test in year t is easier than the test in year t-1, one expects test scores to be relatively higher in the current year making the marginal impacts of a higher score in the current year different from those of the preceding year. Taking the percentage point difference from the average state score means that testing years with low scores resulting from relatively more difficult tests is not going to affect the comparisons of state performance in different time periods. For every testing year, what will matter is how a state scores in comparison to the cross-state average of that year. Secondly, the values of the scores do not mean anything to the average researcher unless they are familiar with the testing and what the typical scores are. The percentage point difference facilitates interpretations of the quality of education measure when going over descriptive statistics since a benchmark (average performance across all states) is imbedded in the measure.

As previously mentioned, this essay focuses on cross-state panel-data investigation. However, a brief cross-state cross-sectional study is done. The variables in the cross-sectional version of the state data are averages for the entire 1990-2007 period.

The panel version of the data contains five periods: 1990-1993, 1994-1997, 1998-2001, 2002-2005, and 2006-2007. All variables in this version of the data have been transformed to period averages corresponding to the five time intervals. Since the analysis deals with macroeconomic variables, I omit 2008 and 2009 as these were years where the U.S. was in the midst of a severe recession. Data points taken from these years would likely not be comparable to past years. As is sometimes done with data during World War II (or atypical periods), simply omitting the years seemed to be the most prudent.

The tables that follow present summary statistics for the cross-sectional version and the panel version of the data.

Table 4a: Summary statistics for the cross-sectional version of the data

Variable	Observations	Mean	Std. Dev.	Minimum	Maximum
inc gr	50	1.62	0.38	0.82	3.00
gini	50	**53.85**	2.05	50.29	58.48
pov	50	10.93	2.83	6.22	17.73
inc1990	50	23.41	3.65	16.72	33.40
white	50	84.95	12.13	27.06	97.72
black	50	9.23	9.21	0.24	35.78
pop gr	50	1.11	0.85	0.01	4.48
ump rate	50	5.08	0.92	2.99	7.08
school	50	12.85	0.34	12.09	13.35
qed	50	0.23	1.52	-3.46	2.73
bach	50	19.23	3.52	12.01	27.50

Table 4b: Summary statistics for the panel version of the data

Variable	Observations[1]	Mean	Std. Dev.	Minimum	Maximum
inc gr	173	1.99	1.07	-1.30	5.74
gini	173	**54.85**	2.63	47.77	60.90
pov	173	11.05	2.88	5.09	19.97
white	173	83.22	13.25	20.70	98.71
black	173	10.26	10.03	0.12	41.82
pop gr	173	1.01	0.83	-1.26	3.90
ump rate	173	4.71	1.08	2.56	7.69
school	173	12.97	0.37	12.01	13.85
qed (lag 1)	173	0.17	1.68	-3.99	3.48
bach	173	21.29	4.36	11.56	36.74

1 - We have 173 observations as some states did not have data for the lagged *qed* for every period.

The reason why the quality of education variable is lagged one period in the panel version of the data is because the mathematics test scores used are for students in grade 8 (i.e. students 13 to 14 years of age). Lagging the variable means that for any given period, the test scores correspond to those approximately 14 to 22 years of age. This allows the variable to better reflect the level of human capital for the current period. Ideally, one would want test scores for every age group in the population for all periods. Unfortunately, this is not possible due to data limitations.

Scores for the students in grade 12 would have been preferred for this analysis yet, because testing in this group occurred infrequently, the grade 8 scores were deemed more useful. In addition, it seems reasonable to think that states whose school systems produce high grade 8 test scores would also produce high grade 12 test scores. Thus, at least loosely, the lagged test scores for a given period can be thought to represent the quality of human capital for those aged 14 to 26 (as those aged 18 at the beginning of the previous four year period would be 26 by the end of the current four year period). To some degree, the test scores could also represent the level of investment in the quality of human capital. Consequently, if the existing quality of human capital is persistent, than an influx of higher quality human capital should, at least theoretically, have measurable macroeconomic impacts.

The quality measure in this paper is by no means perfect, yet it should allow for meaningful results to be obtained from the panel-data analysis, which is the main focus of this study because of the insight that can be gained from the estimation methods that cannot be used on cross-sectional data.

Tables 5a and 5b display correlations between variables for both versions of the data.

Table 5a: Correlation table for the cross-sectional version of the data

(obs=50)	inc gr	gini	pov	inc 1990	white	black	pop gr	ump rate	school	qed	Bach
inc gr	1	0.04	0.26	-0.33	0.28	-0.10	-0.08	-0.26	-0.03	0.12	-0.07
gini	0.04	1	0.65	-0.03	-0.25	0.40	0.13	0.62	-0.54	-0.59	-0.15
pov	0.26	0.65	1	-0.61	-0.19	0.35	0.02	0.49	-0.83	-0.67	-0.64
inc1990	-0.33	-0.03	-0.61	1	-0.17	-0.06	-0.09	0.01	0.56	0.26	0.76
white	0.28	-0.25	-0.19	-0.17	1	-0.58	0.00	-0.13	0.17	0.60	0.00
black	-0.10	0.40	0.35	-0.06	-0.58	1	-0.02	0.23	-0.39	-0.58	-0.11
pop gr	-0.08	0.13	0.02	-0.09	0.00	-0.02	1	0.00	-0.07	-0.22	-0.12
ump rate	-0.26	0.62	0.49	0.01	-0.13	0.23	0.00	1	-0.43	-0.50	-0.24

school	-0.03	-0.54	-0.83	0.56	0.17	-0.39	-0.07	-0.43	1	0.71	0.79
qed	0.12	-0.59	-0.67	0.26	0.60	-0.58	-0.22	-0.50	0.71	1	0.49
bach	-0.07	-0.15	-0.64	0.76	0.00	-0.11	-0.12	-0.24	0.79	0.49	1

Table 5b: Correlation table for the panel version of the data

(obs=174)	inc gr	gini	pov	white	Black	pop gr	ump rate	school	qed	Bach
inc gr	1	-0.05	-0.05	0.12	-0.12	-0.05	-0.47	0.12	0.08	0.10
gini	-0.05	1	0.43	-0.26	0.32	0.06	0.38	-0.13	-0.43	0.16
pov	-0.05	0.43	1	-0.16	0.33	0.03	0.53	-0.75	-0.61	-0.58
white	0.12	-0.26	-0.16	1	-0.58	0.05	-0.11	0.08	0.58	-0.05
black	-0.12	0.32	0.33	-0.58	1	-0.03	0.21	-0.27	-0.57	-0.07
pop gr	-0.05	0.06	0.03	0.05	-0.03	1	-0.08	-0.15	-0.12	-0.14
ump rate	-0.47	0.38	0.53	-0.11	0.21	-0.08	1	-0.43	-0.35	-0.27
school	0.12	-0.13	-0.75	0.08	-0.27	-0.15	-0.43	1	0.62	0.82
qed (lag 1)	0.08	-0.43	-0.61	0.58	-0.57	-0.12	-0.35	0.62	1	0.40
bach	0.10	0.16	-0.58	-0.05	-0.07	-0.14	-0.27	0.82	0.40	1

The correlations between *qed* (quality measure) and the three dependent variables (income per capita rate of growth, income inequality, and rate of poverty) are all as expected; for both datasets, quality is positively correlated with *inc gr* (though not highly), and negatively correlated with *gini* and *pov*. Furthermore, the magnitude of the correlations are more or less as strong or stronger than the correlations between years of schooling (quantity measure) and the dependent variables[14]. These preliminary findings show that in this dataset there seems to be a significant relationship, in the expected directions, between the quality of education measure (which is argued to be a proxy for quality of human capital) and the dependent variables.

In terms of the modeling for the regressional analysis, I assume linear relationships. It is quite possible that the relationships are non-linear however, for simplicity, I use linear models.

[14] The correlation between *school* and *inc gr* for the cross-sectional version of the data has the opposite sign of what theory leads us to expect. However, the correlation is so weak (-0.03) that, at least from a preliminary point of view, it seems more reasonable to think that the variables are unrelated.

Except for growth, there exists little literature on the form of the relationship between my chosen dependent variables and explanatory variables. Future studies could focus on developing proper models that explain the interactions between these variables.

The Results

The expected signs for the regression coefficients are shown in Table 6.

Table 6: Expected sign of regression coefficients

Variable	inc gr	gini	pov
inc1990	negative	n/a	n/a
white	positive	negative	negative
black	negative	positive	positive
pop gr	negative	positive	positive
ump rate	negative	positive	positive
school	positive	negative	negative
qed	positive	negative	negative
bach	positive	indeterminate	indeterminate

The expected signs on *white* and *black* are based on the socioeconomic situations of the two groups. It is no secret that the African American population in the U.S. systematically suffers from higher unemployment rates, lower educational attainment, lower incomes, and generally fare worse than their Caucasian counterparts. The signs for *pop gr*, *ump rate*, *school*, and *qed* need no explanation as they are the same as usual.

The impact of *bach* on *inc gr* is clear. More engineers, lawyers, economists, etc. implies more human capital and, subsequently, a higher growth rate. This argument likely does not hold for developing countries as these countries are in very different stages of development compared to the U.S. states. An influx of highly educated individuals into a poor agricultural country is not expected to have high impacts on growth.

The impact of having a larger (or smaller) percentage of your population with a Bachelor's degree on income inequality and poverty is not clear. Because more educated people tend to command more resources than less educated people, if only a small minority have higher education, then it may be the case that having more educated people would lead to higher income inequality and more poverty. Countering this, there is the impact on growth. If having a greater number of higher educated people leads to higher growth and thus more wealth, then, as was argued earlier, how this extra wealth is distributed would also affect income inequality and the poverty rate. How these forces balance will end up determining the sign of the coefficients.

Cross-Sectional Results

The first results presented are those for the cross-sectional data. For the 1990 to 2007 period, Table 7a displays the results for the regression of average annual growth on initial income per capita and the average values of the rest of the explanatory variables.

Table 7a: Explaining average growth in income per capita - Cross-sectional data (1990-2007), OLS

inc gr	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
inc1990	-0.03*	**-0.11**	-0.03*	**-0.11**	-0.05*	**-0.18**
	[0.02]		[0.02]		[0.03]	
white	0.01**	0.12	0.01**	0.12	0.01*	0.12
	[0.00]		[0.01]		[0.01]	
black	0	0	0	0	0	0
	[0.01]		[0.01]		[0.01]	
pop gr	-0.05	-0.04	-0.07	-0.06	-0.06	-0.05
	[0.04]		[0.05]		[0.06]	
ump rate	-0.1	-0.09	-0.13*	-0.12	-0.12*	-0.11
	[0.07]		[0.07]		[0.07]	
school	0.03	0.01	0.2	0.01	-0.08	-0.03
	[0.20]		[0.28]		[0.44]	
qed			-0.08	-0.12	-0.08	-0.12
			[0.07]		[0.06]	
bach					0.05	0.18

			[0.05]
constant	1.8	-0.53	2.73
	[2.52]	[3.67]	[5.45]
Observations	50	50	50
R-squared	0.23	0.25	0.28

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

Table 7a shows that the only consistently significant explanatory variables are initial income and the fraction of the population that is white. The unemployment rate is also weakly significant in models [2] and [3] where more educational variables are included. For initial income, the negative and weakly significant coefficients found in all three models give reason to believe that the theory of conditional convergence holds in the dataset. The one standard deviation impacts[15] range from a -0.11 to -0.18. The coefficient on unemployment is also negative, as expected, and its one standard deviation impact on growth is around -0.1 percentage points. One thing to notice is that the R-squared for all three models is not particularly high and thus important variables are likely being excluded. Past research using a similar model, but with cross-country data, attained R-squared values of around 0.70[16]. The variables here explain only some of the variance in growth rates across the U.S. states and, most importantly, the variable used for the quality of education is not different from zero in any of the models.

Continuing in the analysis I present, in Table 7b, the cross-sectional results where average inequality is the dependent variable.

[15] One standard deviation impacts are used as they are what a state could expect to achieve by increasing their educational quality.
[16] See Hanushek and Kimko's (2000) baseline estimates in Table 2.

Table 7b: Explaining average inequality in personal income - Cross-sectional data (1990-2007), OLS

gini	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
white	-0.01	-0.12	0	0.00	0.01	0.12
	[0.01]		[0.03]		[0.02]	
black	0.04	0.37	0.03	0.28	0.01	0.09
	[0.02]		[0.03]		[0.02]	
pop gr	0.29	0.25	0.22	0.19	0.27	0.23
	[0.33]		[0.35]		[0.31]	
ump rate	1.03***	0.95	0.95***	0.87	**0.79*** **	0.73
	[0.30]		[0.35]		[0.24]	
school	-1.55*	-0.53	-1.06	-0.36	**-4.14*** **	**-1.41**
	[0.81]		[1.29]		[1.39]	
qed			-0.22	-0.33	-0.28	-0.43
			[0.39]		[0.32]	
bach					0.34***	1.20
					[0.10]	
constant	68.50***		61.76***		95.30***	
	[10.91]		[17.45]		[17.07]	
Observations	50		50		50	
R-squared	0.51		0.52		0.63	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

In terms of explanatory power, the independent variables do much better in explaining income inequality, as measured by the personal income gini coefficient, than growth. Yet, again, quality of education is insignificant in both of the models it appears in. In the third model, where quality and percent with a Bachelor's degree or higher are included, *ump rate*, *school*, and *bach* are all highly statistically significant. The coefficient on unemployment rate implies that an increase of one percentage point in unemployment would increase the gini by 0.79 units. With the mean gini being 53.85, this represents a 1.5% increase in inequality for a state having the mean gini value[17]. For schooling, the coefficient implies that an additional year in average years of schooling would decrease the gini by 4.14 units or approximately 8% for a state with the mean

[17] A gini coefficient of 0 represents perfect equality and a gini of 1 represents maximum concentration.

value. The one standard deviation impact for schooling is less at -1.41. Finally, all else being equal, an increase in the percentage of the population with a Bachelor's degree or more is positively correlated with more inequality.

Table 7c displays the results for the poverty variable.

Table 7c: Explaining average poverty rate - Cross-sectional data (1990-2007), OLS

pov	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
white	-0.01	0.12	0.01	0.12	0.01	0.12
	[0.01]		[0.03]		[0.03]	
black	0	0.00	0	0.00	0	0.00
	[0.03]		[0.03]		[0.04]	
pop gr	-0.1	-0.09	-0.21	-0.18	-0.22	-0.19
	[0.31]		[0.33]		[0.33]	
ump rate	0.48*	0.44	0.36	-0.33	0.38	-0.35
	[0.25]		[0.26]		[0.27]	
school	-6.30***	-2.14	-5.53***	-1.88	-5.22***	**-1.77**
	[0.70]		[0.89]		[1.09]	
qed			-0.34	-0.52	-0.34	-0.52
			[0.40]		[0.42]	
bach					-0.03	-0.11
					[0.12]	
constant	90.20***		79.54***		76.22***	
	[9.86]		[12.87]		[13.53]	
Observations	50		50		50	
R-squared	0.71		0.71		0.71	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

Here, the only consistently significant variable is average years of schooling. Furthermore, the coefficient on quality of education is not different from zero in both models where it is included. The one standard deviation impact for schooling is -1.77 percentage points in model [3]. For a state with the mean level of poverty (10.93%), this translates to a 16.2% drop. This is by no means a small decrease.

The key finding that should be taken away from the cross-sectional analysis is that either test scores are a poor measure of the quality of education and/or the quality of education plays essentially little to no role in determining the levels of the dependent macroeconomic variables.

Panel Results

Continuing the analysis, in Table 8a I present the baseline pooled OLS results for growth in income per capita.

Table 8a: Explaining average growth in income per capita - Panel-data (1990-2007), pooled OLS

inc gr	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
white	0.00	0.00	0.01	0.13	0.01	0.13
	[0.01]		[0.01]		[0.01]	
black	0.00	0.00	0.00	0.00	-0.01	-0.10
	[0.01]		[0.01]		[0.01]	
pop gr	-0.15	-0.12	-0.17*	-0.14	-0.17*	-0.14
	[0.10]		[0.10]		[0.10]	
ump rate	-0.35***	-0.38	-0.37***	-0.40	-0.38***	-0.41
	[0.08]		[0.08]		[0.08]	
school	-0.19	-0.07	0.16	0.06	-0.18	-0.07
	[0.21]		[0.28]		[0.44]	
qed (lag 1)			-0.13**	-0.22	-0.13**	-0.22
			[0.06]		[0.06]	
bach					0.03	0.13
					[0.03]	
Period Dummies[1]						
1998-2001	0.33*		0.26		0.26	
	[0.17]		[0.17]		[0.17]	
2002-2005	-0.79***		-0.84***		-0.83***	
	[0.20]		[0.19]		[0.19]	
2006-2007	0.26		0.13		0.14	
	[0.21]		[0.22]		[0.21]	
constant	5.90*		1.03		4.78	
	[3.29]		[4.16]		[5.57]	
Observations	173		173		173	
R-squared	0.38		0.40		0.40	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

1 - No time dummy for the first period is included and thus it becomes the base case.

Unlike in the cross-sectional analysis, quality of education is statistically different from zero. However, the conditional relationship is in the opposite direction as one would expect and as is found in other studies. At roughly a fall of a quarter of a percentage point, the impact of a one standard deviation increase in average test score is not trivial. This counterintuitive finding is problematic and signifies that the model may suffer from an omitted variable bias. More on this will be discussed later.

The other finding worth mentioning is the large and negative impact that unemployment has on growth. The negative relationship is expected, but the magnitude of the impact may not be. A one percentage point increase in unemployment is related with, approximately, a 0.40 point decrease in the average (period) growth rate.

Table 8b shows the panel results for income inequality.

Table 8b: Explaining average inequality in personal income - Panel-data (1990-2007), pooled OLS

gini	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
white	-0.01	-0.13	0.01	0.13	0.02	0.27
	[0.01]		[0.01]		[0.01]	
black	0.03*	0.30	0.02	0.20	0.00	0.00
	[0.02]		[0.02]		[0.02]	
pop gr	0.44**	0.37	0.37**	0.31	0.36*	0.30
	[0.19]		[0.19]		[0.19]	
ump rate	1.02***	1.10	0.95***	1.03	0.80***	0.86
	[0.17]		[0.16]		[0.13]	
school	-0.99*	-0.37	0.13	0.05	-4.21***	-1.56
	[0.54]		[0.76]		[0.79]	
qed (lag 1)			-0.43***	-0.72	-0.36***	**-0.60**
			[0.16]		[0.12]	
bach					0.38***	1.48
					[0.06]	
Period Dummies[1]						
1998-2001	2.20***		1.95***		1.97***	
	[0.38]		[0.40]		[0.36]	
2002-2005	2.99***		2.81***		3.04***	
	[0.38]		[0.40]		[0.34]	
2006-2007	4.29***		3.89***		4.02***	
	[0.43]		[0.49]		[0.41]	
constant	60.38***		44.80***		93.48***	
	[7.47]		[10.34]		[9.77]	
Observations	173		173		173	
R-squared	0.52		0.55		0.67	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

1 - No time dummy for the first period is included and thus it becomes the base case.

For inequality, average test scores are highly statistically significant and have the expected negative conditional relationship that economic theory would predict. For the third model, the one standard deviation impact is -0.60. This translates to roughly a 1.1% decrease in the gini for a state having a gini equal to the mean value (54.85). Other than the population growth rate, which we now find to be statistically different from zero, the other results are quite comparable to the cross-sectional analysis.

Table 8c: Explaining average poverty rate - Panel-data (1990-2007), pooled OLS

poverty	[1]	One Std. Impact	[2]	One Std. Impact	[3]	One Std. Impact
white	0.00	0.00	0.01	0.13	0.01	0.13
	[0.01]		[0.01]		[0.01]	
black	0.02	0.20	0.01	0.10	0.01	0.10
	[0.02]		[0.02]		[0.02]	
pop gr	-0.16	-0.13	-0.21	-0.17	-0.21	-0.17
	[0.19]		[0.18]		[0.19]	
ump rate	0.64***	0.69	0.60***	0.32	**0.61*** **	0.66
	[0.14]		[0.14]		[0.14]	
school	-5.71***	-2.11	-5.03***	-1.86	-4.85***	-1.79
	[0.42]		[0.48]		[0.67]	
qed (lag 1)			-0.26*	-0.44	-0.27*	-0.45
			[0.15]		[0.15]	
bach					-0.02	-0.09
					[0.06]	
Period Dummies[1]						
1998-2001	0.41		0.26		0.26	
	[0.37]		[0.38]		[0.38]	
2002-2005	0.73*		0.62		0.61	
	[0.41]		[0.41]		[0.41]	
2006-2007	1.84***		1.59***		1.59***	
	[0.40]		[0.39]		[0.39]	
constant	81.35***		71.80***		69.75***	
	[5.99]		[6.84]		[8.18]	
Observations	173		173		173	
R-squared	0.68		0.69		0.69	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

1 - No time dummy for the first period is included and thus it becomes the base case.

I finish off the pooled OLS results with Table 8c which shows what was found for the panel-data regressions explaining the poverty rate. The magnitude of schooling's impact is roughly the same as in the cross-sectional regression. The coefficient on *ump rate* is statistically different from zero in all three models and has the expected sign. For the third model, the conditional relationship implies that a one percentage point increase in the rate of unemployment is associated with a 0.61 percentage point increase in the poverty rate. For a state with the mean poverty rate this represents a 5.6% increase. The coefficient on the variable of interest, the lagged value of *qed*, is statistically different from zero at the 10% level of significance. It's one standard deviation impact, at roughly 0.45, is non-trivial but smaller than the impact associated with average years of schooling. As well, the addition of quality of education does decrease the impact associated with the quantity of education.

The baseline pooled OLS results show that test scores share a significant relationship with the macroeconomic variables studied. However, for the growth model, a negative instead of a positive conditional relationship is found. In reality, it is highly unlikely that higher test scores would lead to lower growth. What is likely being observed is a correlation between test scores and some omitted variable. The fixed and random effects models presented later will shed more light on this peculiar finding.

The results from the cross-sectional OLS and the pooled OLS differ significantly with respect to the variable of interest. This finding gives reason to believe that the dynamics of the relationships between the variables, which are washed out in the cross-sectional analysis, are important. This is not surprising as the U.S. states are not homogeneous entities. There is no reason to think that the effects of changes in educational quality in New York are going to be exactly the same as those in Nebraska. Like in many other studies, the evolution of the variable of interest (quality) within units, states in this case, is likely more informative than the changes between units.

Another interesting finding that differs from past research, although not directly as different dependent variables are considered, is that the one standard deviation impact of quantity of schooling is larger in every case than the quality of schooling. This finding, assuming the correct causal relationship holds, implies that if the cost to a government of slightly increasing quantity and quality are the same, then state governments may want to target quantity as its impact on income inequality and poverty are roughly three times as great.

To assess whether or not *qed* is exogenous in the models and can be included as a right hand side variable, I perform Granger causality tests. If the dependent variable Y *(inc gr*, *gini*, or *pov*) is said to Granger cause *qed*, than there is statistical evidence that implies *qed* is endogenous. Table 9 shows the result of these Granger causality tests.

Table 9: Granger causality testing - does Y (inc gr, gini, pov) granger cause X (qed)?

Explanatory Variable (X):	qed (Y = inc gr)		qed (Y = gini)		qed (Y = pov)	
Regression:	X on X_{t-i}	X on X_{t-I}, Y_{t-i}	X on X_{t-i}	X on X_{t-I}, Y_{t-i}	X on X_{t-i}	X on X_{t-I}, Y_{t-i}
qed (lag 1)	1.19***	1.14***	1.19***	1.18***	1.19***	1.19***
	[0.17]	[0.14]	[0.17]	[0.19]	[0.17]	[0.17]
qed (lag 2)	-0.13	-0.08	-0.13	-0.1	-0.13	-0.16
	[0.13]	[0.13]	[0.13]	[0.16]	[0.13]	[0.14]
qed (lag 3)	-0.07	0.04	-0.07	-0.16	-0.07	0.00
	[0.19]	[0.21]	[0.19]	[0.25]	[0.19]	[0.23]
qed (lag 4)	0.00	-0.16	0.00	0.05	0.00	-0.06
	[0.15]	[0.18]	[0.15]	[0.18]	[0.15]	[0.18]
Y (lag 1)		0.04		0.05		-0.07
		[0.15]		[0.08]		[0.12]
Y (lag 2)		0.25		-0.03		0.03
		[0.17]		[0.11]		[0.12]
Y (lag 3)		-0.02		-0.04		-0.05
		[0.13]		[0.12]		[0.09]
Y (lag 4)		0.03		0.00		0.07

		[0.10]		[0.09]		[0.07]
constant	0.02	-0.66*	0.02	0.95	0.02	0.24
	[0.09]	[0.36]	[0.09]	[3.66]	[0.09]	[0.52]
Observations	32	32	32	32	32	32
R-squared	0.94	0.95	0.94	0.94	0.94	0.94
H_0: all coefficients on Y_{t-i} are zero						
F-test statistic p-value		0.32		0.89		0.88

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

In this table i=1, 2, 3, 4

Since we fail to reject the null hypothesis (H_0: all coefficients on the lagged dependent variables are jointly zero) and conclude that Y does not Granger cause *qed* in all three cases, there is statistical evidence that implies that *qed* is exogenous in the models and can safely be included as an explanatory variable.

As Granger causality is subject to sampling variability, it is quite possible to find that X Granger causes Y and that Y Granger causes X, or neither Granger causes the other (or some combination of the two). For sake of completeness, I have included the results of the Granger causality tests running the other way in the appendix (see Table A6).

Continuing the study, I now present the fixed and random effects estimates of the coefficients. Both of these estimation methods allow one to take advantage of the time dimension of the data to control for unobservable state specific factors otherwise ignored in cross-sectional analysis. As a refresher, I quickly go over the key points of these estimation methods. Assume that the model can be written as,

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + u_{it}, where\ u_{it} = a_i + e_{it}$$

The subscript *i* is the individual and *t* the time period. The term a_i is constant across time but differs between states.

The benefit of the fixed effects approach is that if the error term in the model can be broken into a state specific constant term a_i and a random idiosyncratic error v_{it} uncorrelated with the explanatory variables, then, given certain assumptions[18], the fixed effects estimators will be unbiased and consistent whereas those for the pooled OLS will not. Furthermore, under homoskedasticity,

$$var(u_{it}|\boldsymbol{X}_i, a_i) = var(u_{it}) = \delta_u^2, for\ all\ t = 1, \ldots, T$$

and unserially correlated errors,

$$cov(u_{it}, u_{is}|\boldsymbol{X}_i, a_i) = 0, for\ t \neq s$$

the fixed effects estimators are, asymptotically[19], the best linear unbiased estimators (BLUE).

The fixed effects approach is one way to correct for an omitted variable problem where the omitted variable, in this case a_i, is constant across time within states.

In cases where there is reason to believe that the unobservable term a_i is uncorrelated with the explanatory variables,

$$cov(x_{jit}, a_i) = 0, for\ t = 1, \ldots, T\ and\ j = 1, \ldots, k$$

one can make use of the random effects method. Under the random effects assumption, pooled OLS also yields consistent estimates but is usually less efficient.

The fixed effects and random effects estimates are presented in Table 10.

Table 10: Random Effects (RE) VS. Fixed Effects (FE)

Dependent Variable:	inc gr		gini		pov	
	RE	FE	RE	FE	RE	FE
white	0.01	-0.18***	0.00	-0.05	0.00	-0.02
	[0.02]	[0.07]	[0.02]	[0.07]	[0.03]	[0.09]
black	-0.01	-0.14*	0.02	0.02	0.05	0.13

[18] The required assumptions can be found in the appendix (see Table A7).
[19] The asymptotic results in the fixed effects case are when T (number of time periods) is small and N→∞ (number of individuals, states for us).

	[0.02]	[0.07]	[0.03]	[0.09]	[0.03]	[0.10]
pop gr	-0.17	-0.28	0.01	-0.45*	-0.04	0.18
	[0.12]	[0.26]	[0.23]	[0.26]	[0.25]	[0.29]
ump rate	-0.45***	**-0.70*****	0.51***	0.30**	0.63***	0.61***
	[0.09]	[0.11]	[0.13]	[0.13]	[0.12]	[0.13]
school	-0.19	0.83	-4.53***	-3.13**	-3.84***	-3.11***
	[0.48]	[0.95]	[0.92]	[1.31]	[0.99]	[1.06]
qed (lag 1)	-0.12*	0.04	-0.14	0.09	-0.13	-0.12
	[0.07]	[0.12]	[0.12]	[0.17]	[0.11]	[0.14]
bach	0.03	-0.04	0.39***	0.45***	-0.01	0.12
	[0.04]	[0.10]	[0.08]	[0.12]	[0.09]	[0.10]
Period Dummies						
1998-2001	0.19	-0.25	1.71***	0.99***	0.00	-0.43
	[0.17]	[0.31]	[0.26]	[0.32]	[0.24]	[0.28]
2002-2005	-0.83***	-1.57***	2.81***	1.75***	0.06	-0.82**
	[0.19]	[0.52]	[0.28]	[0.46]	[0.31]	[0.41]
2006-2007	0.09	-0.98*	3.78***	2.52***	1.18***	0.12
	[0.21]	[0.59]	[0.32]	[0.55]	[0.31]	[0.44]
constant	5.67	13.26	100.61***	87.90***	56.98***	46.09***
	[6.10]	[12.56]	[11.24]	[16.90]	[12.07]	[14.46]
observations	173	173	173	173	173	173
rho	0.18	0.89	0.67	0.84	0.8	0.88
Hausman Test						
Chi^2 Statistic	36.38		37.76		17.04	
H_0: Both FE and RE yield consistent estimates; Ha: RE yields inconsistent estimates, FE yields consistent estimates.						
P-value	0.00		0.00		0.07	

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets. No time dummy for the first period is included and thus it becomes the base case.

Comparing the fixed effects estimates with the random effects estimates shows that, for the most part, they are similar in terms of their statistical significance and, in some cases, magnitude. In order to ascertain which estimates are more reliable, a Hausman test is performed and the test statistics are reported at the bottom of Table 10. For all three dependent variables, the test results give reason to believe that the random effects estimates are inconsistent and we should then focus on the fixed effects regressions.

Comparing the fixed effects estimates for the coefficients with the equivalent pooled OLS values shows some differences. For growth in income per capita, the coefficient on our variable of interest, *qed (lag 1)*, is no longer statistically different from zero. The coefficient on *ump rate* is still statistically different from zero and has almost doubled in magnitude (-0.70 instead of -0.38). A peculiar finding in the fixed effects model is that the coefficient on *white* has a negative and significant coefficient.

As in the model for *inc gr*, in the income inequality model the quality of education is no longer statistically significant. Another finding in the fixed effects regression for income inequality is that the sign of the coefficient on *pop gr* has switched to negative from being positive in the pooled OLS model, though the coefficient is only weakly different from zero. The coefficient on the quality of education in the fixed effects model for poverty is also insignificant.

In summary, once state fixed effects are controlled for, the quality of education seems to play no role in determining any of the dependent variables. Thus, there is reason to question the ability of test scores to proxy the quality of human capital.

Because I have chosen to use robust standard errors, heteroskedasticity is not a major concern in my models. However, for unbiased and consistent estimates of the standard errors, we need to assume no serial correlation across the idiosyncratic errors. If the models do suffer from serial correlation, then the standard errors will be biased and the estimated coefficients will be less efficient. Fortunately, a test for the presence of serial correlation in fixed effects models does exist. Table 11 shows the results of the Wooldridge test for serial correlation in fixed effects models[20].

[20] This test is quite useful as it is still valid (although perhaps less so) in the presence of heteroskedasticity and unbalanced panel-data (Drukker 2003).

Table 11: Wooldridge test for serial correlation in panel-data fixed effects model[1]

H_0: no first-order serial correlation	Dependent Variable		
	inc gr	gini	Pov
F-statistic	2.44	2.44	15.76
P-value	0.13	0.13	0.00
Observations[2]	109	109	109

1 - as described in Drukker (2003).

2 - We now have 109 observations instead of 173 as the Wooldridge test uses first differencing to remove the state specific fixed effects. Because of the first differencing we lose some observations.

The results of the serial correlation test show that we fail to reject the null hypothesis of no serial correlation for two of the three estimated models. However, for *inc gr* and *gini*, at a slightly lower level of statistical significance (α = 0.14 for example) we would reject H_0, meaning serial correlation could be a problem for these two fixed effects models as well.

To address the possible issue of serial correlation, I re-estimate the coefficients using the generalised method of moments (GMM) procedure developed by Arellano and Bond (1991). This procedure is summarized below.

Consider a standard linear panel-data model with individual specific fixed effects and serially correlated errors as follows.

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + a_i + u_{it},\ where\ u_{it} = \rho u_{it-1} + e_{it}\ \ and\ \ e_{it} \sim N(0, \delta^2)\ [1]$$

Here ρ is a parameter and e_{it} is considered to be white noise. First differencing with period t-1 gives,

$$\Delta y_{it} = \beta_1 \Delta x_{1it} + \cdots + \beta_k \Delta x_{kit} + \Delta u_{it},\ where\ \Delta u_{it} = \rho \Delta u_{it-1} + \Delta e_{it}\ [2]$$

After a little manipulation, it can be shown that,

$$\Delta u_{it-1} = \Delta y_{it-1} - \beta_1 \Delta x_{1it-1} - \cdots - \beta_k \Delta x_{kit-1}\ [3]$$

Using [3] we get,

$$\Delta y_{it} = \beta_1 \Delta x_{1it} + \cdots + \beta_k \Delta x_{kit} + \rho \Delta y_{it-1} - \rho \beta_1 \Delta x_{1it-1} - \cdots - \rho \beta_k \Delta x_{kit-1} + \Delta e_{it}\ \ [4]$$

This transformation addresses the issues of individual specific fixed effects and an error term following and AR(1) process of autocorrelation. To estimate this corrected model one must use an instrument for Δy_{it-1} since consistency depends on the following condition holding.

$$cov(\Delta y_{it-1},\ \Delta e_{it}) = 0$$

This condition is unlikely to hold as the exogenous shock e_{it-1} is most likely correlated with y_{it-1}. For example, an increase in the growth rate due to some unobservable shock would manifest itself in our model by way of the error term. Expanding the covariance above gives,

$$\begin{aligned} cov(\Delta y_{it-1}, \Delta e_{it}) &= cov(y_{it-1} - y_{it-2}, e_{it} - e_{it-1}) \\ &= E[y_{it-1}e_{it}] - E[y_{it-1}e_{it-1}] - E[y_{it-2}e_{it}] + E[y_{it-2}e_{it-1}] \\ &\neq 0 \ as\ E[y_{it-1}e_{it-1}] \neq 0 \end{aligned}$$

Following the Arellano-Bond procedure, we can instrument for Δy_{it-1} with y_{it-2}, x_{1it}, ..., and x_{kit}. The Arellano-Bond[21] results are shown in Table 12.

Table 12: Arellano-Bond Difference (two-step) GMM estimates

Dependent Variable (dep var):	inc gr	gini	Pov
LD[1] dep var[2]	0.10	-0.67	0.06
	[0.27]	[0.59]	[0.15]
D[3] white	-0.11	-0.09	0.05
	[0.11]	[0.08]	[0.08]
LD white	0.06	-0.21*	0.12
	[0.08]	[0.12]	[0.09]
D black	0.02	-0.1	0.36***
	[0.10]	[0.18]	[0.08]
LD black	-0.13	-0.12	0.04
	[0.08]	[0.15]	[0.08]
D pop gr	-0.46*	-0.91***	-0.16
	[0.26]	[0.32]	[0.25]
LD pop gr	-0.41	-0.12	-0.13
	[0.30]	[0.39]	[0.26]

[21] A more detailed explanation of the Arellano-Bond and other different GMM procedures and how to run them in Stata can be found in Roodman (2006). Another useful source which explains the theory behind GMM estimators and dynamic panel-data models can be found in Bond (2002).

D ump rate	-0.61***	0.15	0.02
	[0.18]	[0.18]	[0.16]
LD ump rate	0.08	0.57	0.18
	[0.22]	[0.39]	[0.18]
D school	-0.15	-4.86**	-2.93
	[1.73]	[2.30]	[2.00]
LD school	2.16*	-3.66	-0.03
	[1.29]	[2.49]	[1.48]
D qed (lag 1)	**0.54*** **	-0.21	-0.02
	[0.13]	[0.31]	[0.18]
LD qed (lag 1)	0.17	0.58***	0.41***
	[0.12]	[0.19]	[0.15]
D bach	0.11	0.51***	0.24
	[0.10]	[0.18]	[0.15]
LD bach	0.00	0.38*	0.27**
	[0.11]	[0.21]	[0.12]
Period Dummies			
D 2002-2005	-1.85***	2.29**	-0.27
	[0.55]	[1.16]	[0.47]
D 2006-2007	-0.93	2.67	-0.23
	[0.64]	[1.72]	[0.55]
Observations[4]	65	65	65
Number of Groups	33	33	33

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

1-LD refers to lagged first difference.

2-Instrumented for using lagged levels (beginning at t-2) of the dependent and independent variables.

3-D refers to first difference.

4-Because the Arellano-Bond procedure requires that we difference the variables I lose the two earliest cross-sections (as qed is already lagged one period in the original model).

For income growth, the Arellano-Bond estimate for the coefficient on educational quality, while controlling for quantity, is 0.54 and is now statistically significant with the appropriate positive sign. A one percentage point increase in *D qed (lag 1)* would increase *inc gr* by 0.54 points according to this new estimate. The one standard deviation impact on growth is 0.37[22] percentage points. However, the Arellano-Bond estimates require valid instrumental

[22] The standard deviation of *qed (lag 1)* in difference form is 0.69 for the 65 observations.

variables to be viable. Table 13 gives reason to believe that the Arellano-Bond transformation corrected any issue of serial correlation and shows that we fail to reject the null hypothesis of valid instruments using the Hansen-Sargan test for over-identifying restrictions. Therefore, the Arellano-Bond results for the income growth model may be seen as more reliable than the fixed effects ones shown earlier.

Table 13: Testing for serial correlation and over-identifying restrictions in Difference GMM model

Dependent Variable:	inc gr	gini	Pov
	Arellano-Bond test for AR(1) in first-differenced residual H_0: no AR(1)		
Normal: P-value	0.99	0.33	0.93
	Hansen-Sargan test for over-identifying restrictions: validity of moment conditions H_0: moment conditions hold, instruments are valid		
Chi^2: P-value	0.25	0.05	0.12

Note: We need not check for AR(2) as there are only two periods in the first difference estimation. Also note that tests statistics are based on asymptotic properties.

For the remaining macroeconomic variables, income inequality and poverty rate, correcting for autocorrelation in the error term to attain unbiased standard errors did not change the statistical significance of the coefficients on the variable of interest. The coefficients on *D qed (lag 1)* for the *gini* and *pov* models are statistically not different from zero. This finding reaffirms the results found earlier; an increase in *qed (lag 1)* has essentially no impact on inequality or poverty. In terms of the validity of the instruments, the Hansen-Sargan test reveals that the instruments in the *gini* model are likely not valid, and the ones for *pov* may not be either as the p-value is only 0.12. This gives evidence to question the Arellano-Bond estimates for these two dependent variables.

One problem with the Arellano-Bond Difference GMM procedure is that the reliability of the estimates depends on the strength of the instruments; weak instruments give weak or even invalid results. To try and get around this, one can make use of the Arellano-Bover/Blundell-

Bond System GMM procedure. Like the Difference GMM method, this method assumes that strong instrumental variables are unavailable and so researchers must make do with the data at hand. The advantage this procedure has over Difference GMM is that, given certain additional conditions holding, the transformation used gives rise to further possible instruments for the endogenous lagged dependent variable. I briefly describe the fundamental idea behind the procedure[23].

Equation [1] can be written as,

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + a_i + u_{it} \text{ [5]}$$

The variables are the same as before. Notice that,

$$u_{it-1} = y_{it-1} - \beta_1 x_{1it-1} - \cdots - \beta_k x_{kit-1} - a_i$$

Using this we can get,

$$y_{it} = (1-\rho)\beta_0 + \rho y_{it-1} + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} - \rho\beta_{k1} x_{1it} - \cdots - \rho\beta_k x_{kit}$$

$$+(1-\rho)a_i + e_{it} \text{ [6]}$$

Now, instead of instrumenting for differences with lagged levels for the difference equation, we instrument for levels with lagged differences in the levels equation. If we assume the series for the dependent variable is convergent, it can be shown that,

$$E[\Delta y_{it-1} u_{it}] = 0,$$

$$as\ E[\Delta y_{it-1} a_i] = 0 \text{ [7]}$$

Therefore we can instrument for the endogenous variable $y_{it-1}$with Δy_{it-1}. Furthermore, in addition to any exogenous instruments, which need not be differenced or lagged to qualify as valid instruments, for predetermined explanatory variables one can show that the following moment condition will hold.

[23] See Roodman (2006) and Bond (2002) for a more detailed explanation of System GMM.

$$E[\Delta x_{kit} u_{it}] = 0$$

$$if\ E[\Delta x_{kit} a_i] = 0\ [8]$$

This means that Δx_{kit} is also a valid instrument. Another advantage that System GMM has over Difference GMM is that the transformation used allows us to lose only one period instead of two. When the data only has a few periods this can be very attractive.

One must note that if we believe conditions [7] and [8] do not hold, the System GMM estimates are not valid. Given that the coefficients on quality of education became statistically insignificant once the state fixed effects were differenced out, as mentioned earlier, there is evidence to believe that quality is correlated with and is picking up some of the impact of the state fixed effects. Because of this, there is more than a small possibility that conditions [7] and [8] do not hold in my data. Keeping this in mind, I present the System GMM results in Table 14.

Table 14: Arellano-Bover/Blundell-Bond (two-step) System GMM estimates

Dependent Variable (dep var):	inc gr	gini	pov
L^1 dep var^2	-0.12	0.76	0.33
	[0.31]	[0.53]	[0.21]
White	-0.06	0.17	-0.02
	[0.10]	[0.27]	[0.18]
L white	0.03	-0.32	-0.11
	[0.12]	[0.38]	[0.19]
Black	0.08	0.34	0.40**
	[0.09]	[0.32]	[0.18]
L black	-0.14	-0.41	-0.46**
	[0.10]	[0.44]	[0.20]
pop gr	-0.43	-0.25	-0.89*
	[0.54]	[0.43]	[0.44]
L pop gr	0.02	-0.24	-0.01
	[0.54]	[1.56]	[0.90]
ump rate	-0.75*	0.04	-0.11
	[0.40]	[0.38]	[0.44]
L ump rate	0.11	-0.44	0.4
	[0.25]	[1.20]	[0.71]
School	-6.14	-4.26	-5.6

	[4.71]	[5.65]	[6.25]
L school	2.79	0.05	-0.44
	[2.34]	[3.13]	[1.82]
qed (lag 1)	0.19	0.18	0.47
	[0.28]	[0.53]	[0.32]
L qed (lag 1)	-0.04	0.47	0.64*
	[0.23]	[0.45]	[0.37]
Bach	0.22	0.23	-0.07
	[0.19]	[0.25]	[0.26]
L bach	-0.08	-0.05	0.06
	[0.24]	[0.39]	[0.21]
Period Dummies[3]			
1998-2001	49.34	80.52	95.28
	[44.01]	[88.25]	[79.96]
2002-2005	48.6	81.2	96.96
	[43.78]	[88.50]	[79.48]
2006-2007	49.51	80.8	97.46
	[44.23]	[89.21]	[80.52]
Observations	109	109	109
Number of Groups	44	44	44
Number of Instruments[4]	27	27	27

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets (with Windmeijer's finite-sample correction for the two-step covariance matrix). Also note that, because the sample size is small relative to the total number of instruments available, I collapse the instrument matrix and restrict the GMM type instruments to one and two period lags.

1 - L refers to lagged one period.

2 - Instrumented:

Lagged dependent variables treated as endogenous in the model.

Independent variables treated as predetermined in the model.

Period dummies treated as exogenous in the model

3 - To allow for straight forward comparisons of the coefficients on the period dummies with those in the Arellano-Bond model I omit the constant and include all dummies, removing the need for a base case.

4 - Instruments:

Instruments for first difference equation

GMM type instruments: first and second lags of *L dep var*, *white*, *black*, *pop gr*, *ump rate*, *school*, *qed (lag 1)*, and *bach*.

Instruments for levels equation

Standard instrumental variables: all three period dummies (*1998-2001*, *2002-2005*, and *2006-2007*)

GMM type: first difference of *L dep va*r, *white*, *black*, *pop gr*, *ump rate*, *school*, *qed (lag 1)*, and *bach*.

The System GMM estimates in the income growth model show *qed (lag1)* as being statistically not different from zero, which differs from the Difference GMM coefficient. For the other two dependent variables, the results are the same, educational quality (as measured by test scores) has no impact on income inequality or poverty.

Table 15 displays the results of the serial correlation and over-identifying restrictions tests.

Table 15: Testing for serial correlation and over-identifying restrictions in System GMM model

Dependent Variable:	inc gr	gini	Pov
	Arellano-Bond test for AR(1) in first differenced residual H_0: no AR(1)		
Normal: P-value	0.51	0.03	0.71
	Sargan test for over-identifying restrictions[1]: validity of moment conditions H_0: moment conditions hold, instruments are valid		
Chi^2: P-value	0.00	0.20	0.06
	Hansen test for over-identifying restrictions[2]: validity of moment conditions H_0: moment conditions hold, instruments are valid		
Chi^2: P-value	0.24	0.23	0.60
	Difference-in-Hansen tests of exogeneity of instrument subsets		
GMM instruments for levels equation			
	Hansen test excluding group		
Chi^2: P-value	0.54	0.70	0.36
	Difference H_0: GMM differenced instruments are exogenous		
Chi^2: P-value	0.20	0.17	0.59
Standard instrumental variables			
	Hansen test excluding group		
Chi^2: P-value	0.18	0.25	0.66
	Difference H_0: GMM instruments without standard instrumental variable instruments are exogenous		
Chi^2: P-value	0.47	0.27	0.35

Note: We need not check for AR(2) as there are only two periods in the first difference estimation. Also note that tests statistics are based on asymptotic properties.

1 - The Sargan test in not a robust test, yet it is not weakened when many instruments are used.

2 - The Hansen test is a robust test, yet it is weakened when many instruments are used.

The Arellano-Bond test for AR(1) process in the first differenced residuals shows that the *gini* model may suffer from serial correlation (thus giving biased results). However, this model is the only one where the Sargan test and the Hansen test do not produce contradicting results (both fail to reject the null hypothesis of valid instruments). For the *inc gr* and *pov* System GMM models, the validity of the instruments is more difficult to ascertain. If we had very high p-values for the Hansen test or almost as many instruments as observations one could choose to heed more attention to the less robust Sargan test. Yet, as neither circumstance is true in this instance (since there are 27 instruments versus 109 observations and the highest p-value is 0.60), the robust Hansen test could be thought to be more truthful implying we fail to reject the null hypothesis for these models as well.

Test results for the difference-in-Hansen tests show that, in all cases for all three dependent variables, we fail to reject the null that the GMM instruments are in fact exogenous.

With respect to explaining income inequality or rate of poverty, the conclusions drawn from Difference GMM and System GMM are the same; *qed (lag 1)* plays no role in determining either. In explaining state growth in income per capita, because of the evidence against conditions [7] and [8], I find the Arellano-Bond Difference GMM estimates to be more on track.

In summary, as is found in cross-country analyses, higher test scores seem to indicate higher growth in income per capita. However, this result is not robust, as only when possible serial correlation (AR(1) process) and resulting endogeneity issues are corrected for did we reach this finding. With respect to income inequality and the rate of poverty, once state specific fixed effects are accounted for, test scores seem to become irrelevant.

Conclusion

This study expanded on the recent literature concerning the importance of the quality of education by applying the standardized test score approach on a previously unused panel dataset. In the cross-sectional analysis across the fifty U.S. states for the period of 1990-2007, while controlling for the quantity of education, the quality of education had essentially no impact on the macroeconomic dependent variables.

The appeal of this state dataset over other cross-country datasets was that, with richer testing data, it allowed for a panel-data investigation of the quality of education issue. Controlling for the quantity of education and other macro level factors, the "naive" pooled OLS estimates indicated that the quality of education was a statistically significant factor in the determination of all three macroeconomic dependent variables[24]. Yet, more advanced and more reliable fixed effects estimation, which accounts for the existence of time unvarying state specific factors, provided evidence that the quality of education played no role in the determination of state level per capita income growth, income inequality, or rate of poverty.

The Arellano-Bond Difference GMM procedure that corrects for serial correlation (which testing revealed was likely an issue) also showed that quality of education was not a relevant factor impacting inequality or rate of poverty. This finding did not change when System GMM was applied to address the concern of weak instruments. However, the results from the System GMM method must be taken with caution as there is reason to believe that at least one of the assumptions required for this procedure does not hold.

For per capita income growth, the more sophisticated Arellano-Bond Difference GMM procedure yielded statistically significant positive estimates for the effects of the quality of education on income per capita growth. Additionally, the Sargan test did not reject the null hypothesis of valid instruments. The estimates for the coefficient on educational quality in the income growth model went from statistically significant with the (theoretically speaking) wrong sign using pooled OLS, to statistically insignificant using fixed effects estimation, and finally to

[24] Recall though that the coefficient on *qed (lag 1)* in the pooled OLS income growth model was negative, the opposite of what was expected.

statistically significant with the correct sign using the Arellano-Bond procedure. These inconsistent findings show that the estimation of the impacts of quality is quite susceptible to the estimation procedure chosen.

Although I do not find as clear-cut findings as those of past research, the estimation done in this paper does bring forward evidence that, at least in the U.S., the quality of education may simply be correlated with time unvarying state specific factors. Once these fixed factors are accounted for, the impacts of quality reduce to (statistically) zero for two of the three macroeconomic variables investigated, state level income inequality and rate of poverty.

Keeping the modeling and data limitations in mind, two main conclusions can be drawn from this analysis: (i) state governments may be able to increase the growth rate in income per capita by increasing the quality of education, and (ii) once we control for state specific time unvarying impacts, there is little evidence in this essay that implies that better educational quality would also reduce state level income inequality and/or the rate of poverty.

Bibliography

Barro, Robert J. “Human Capital and Growth.” *The American Economic Review* 91.2 (2001): 12-17. Print.

Barro, Robert J. and JW Lee. “International data on educational attainment: updates and Implications.” *Oxford Economic Papers* 53.3 (2001): 541-63. Print.

Barro, Robert J. and X Sala-i-Martin. *Economic growth, 2nd ed*. Cambridge (MA): MIT Press, 2004. Print.

Bond, Stephen. “Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice.” *Centre for Microdata methods and practice* (2002): Working Paper CWP09/02.

Bond, Stephen and Manuel Arellano. “Some Tests of Specification for Panel-data: Monte Carlo Evidence and an Application to Employment Equations.” *Review of Economic Studies* 58 (1991): 277-297. Print.

Coulombe, Serge and Jean-François Tremblay. “Literacy and Growth.” *Topics in Macroeconomics* 6.2 (2006): Art. 4. Print.

de la Fuente, A. and R. Doménech. “Human capital in growth regressions: how much difference does data quality make? An update and further results. Manuscript”. Instituto de Análisis Económico (CSIC) and Universidad de Valencia (2002).

Drukker, M. David. “Testing for serial correlation in linear panel-data models.” *The Stata Journal* 3.2 (2003): 168-177. Print.

Hanushek, Eric A., and Dennis D. Kimko. “Schooling, Labor-Force Quality, and the Growth of Nations.”*The American Economic Review* 82.5 (2000): 1184-1208. Print.

Hanushek, Eric A. and Dongwook Kim. “Schooling, Labor Force Quality, and Economic Growth.” *National Bureau of Economic Research* (1995): Working Paper 5399.

Hanushek, Eric. *The High Cost of Low Educational Performance: The Long-run Economic Impact of Improving PISA Outcomes*. Program for International Student Assessment (OECD), 2010. Print.

Hanushek, Eric and Ludger Wobmann. *Education Quality and Economic Growth.* Washington DC: The World Bank, 2007. Print.

International Association for the Evaluation of Educational Achievement. *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. United States: Boston College, 1996. Print.

Jamison, Eliot A., Dean T. Jamison and Eric A. Hanushek. "The effects of education and quality on income growth and mortality decline." *Economics of Education Review* 26.6 (2007): 772-789. Print.

"NAEP Overview." *http://nces.ed.gov/nationsreportcard/about/*. National Center for Education Statistics: National Assessment of Educational Progress. 16 August. 2010.

Roodman, David. "How to xtabond2: An Introduction to "Difference" and "System" GMM in Stata." Center for Global Development (2006): Working paper number 103.

Appendix

Table A1: Cross-country cross-sectional analysis - Variables, descriptions, and sources

Variable: 1995-2007 period averages	Description	Source
growth in GDP per capita	The rate of growth in GDP per capita as a percentage	World Bank - WDI[1]
population growth rate	The rate of growth of the population as a percentage	World Bank - WDI
GDP per capita in 1995	Initial GDP per capita at the beginning of the period (1995) in thousands of US$, (base year 2000=100)	World Bank - WDI
years of schooling	(Barro-Lee dataset) Years of schooling by country	World Bank - WDI
quality of education	Quality of education is the international mathematics test scores for grade 8 students scaled to have an average of fifty (as in Hanushek and Kim 1995)	World Bank - WDI

1 - WDI refers to the World Development Indicators dataset compiled by the World Bank.

Table A2: Cross-country cross-sectional analysis - Summary statistics

Variable	Observations	Mean	Std. Dev.	Minimum	Maximum
growth in GDP per capita	62	2.83	1.51	0.85	8.71
population growth rate	62	0.86	1.00	-1.30	3.19
GDP per capita in 1995	62	10.39	9.96	0.23	35.44
years of schooling	62	8.96	1.98	3.91	12.45
quality of education	62	49.57	6.38	25.87	60.01

Table A3: Cross-country cross-sectional analysis - Correlation table

(obs=62)	growth in GDP per capita	population growth rate	GDP per capita in 1995	years of schooling	quality of education
growth in GDP per capita	1	-0.48	-0.44	0.17	0.20
population growth rate	-0.48	1	0.04	-0.44	-0.45
GDP per capita in 1995	-0.44	0.04	1	0.48	0.39
years of schooling	0.17	-0.44	0.48	1	0.57
quality of education	0.20	-0.45	0.39	0.57	1

Table A4: Cross-country cross-sectional analysis - List of countries included in the regression

1. Algeria	21. Ghana	41. Netherlands	61. Ukraine
2. Armenia	22. Greece	42. New Zealand	62. United States
3. Australia	23. Hong Kong SAR, China	43. Norway	
4. Austria	24. Hungary	44. Philippines	
5. Bahrain	25. Iceland	45. Portugal	
6. Belgium	26. Indonesia	46. Romania	
7. Botswana	27. Iran, Islamic Rep.	47. Russian Federation	
8. Bulgaria	28. Ireland	48. Saudi Arabia	
9. Canada	29. Israel	49. Serbia	
10. Chile	30. Italy	50. Singapore	
11. Colombia	31. Japan	51. Slovak Republic	
12. Cyprus	32. Jordan	52. Slovenia	
13. Czech Republic	33. Korea, Rep.	53. South Africa	
14. Denmark	34. Kuwait	54. Spain	
15. Egypt, Arab Rep.	35. Latvia	55. Sweden	
16. El Salvador	36. Lithuania	56. Switzerland	
17. Estonia	37. Malaysia	57. Syrian Arab Republic	
18. Finland	38. Malta	58. Thailand	
19. France	39. Moldova	59. Tunisia	
20. Germany	40. Morocco	60. Turkey	

Table A5: Description of variables attained from the CPS dataset

Age of individuals included in the analysis: 15+	
gini	Constructed using total personal income ("inctot") where all values where deflated to have a base of 1999.
pov	Portion of sample who lived below the poverty line.
white	Portion of sample who identified as being white.
black	Portion of sample who identified as being black.
school	Imputed from the variable "educ". (e.g. high school diploma was taken to imply 12 years of schooling)
bach	Portion of sample with a bachelor's degree or higher.

Note: The state identifier was used to create the variables at the individual state level.

Table A6: Granger causality testing - does X (qed) granger cause Y (inc gr, gini, pov)?

Dependent Variable (Y):	inc gr		gini on lag gini		pov on lag pov	
Regression:	Y on Y_{t-i}	Y on Y_{t-i}, X_{t-i}	Y on Y_{t-i}	Y on Y_{t-i}, X_{t-i}	Y on Y_{t-i}	Y on Y_{t-i}, X_{t-i}
Y (lag 1)	0.82**	0.81**	0.79***	0.78***	0.58***	0.51***
	[0.40]	[0.36]	[0.23]	[0.27]	[0.13]	[0.15]
Y (lag 2)	0.30	0.27	0.33	0.39	-0.06	0.00
	[0.39]	[0.43]	[0.32]	[0.38]	[0.15]	[0.15]
Y (lag 3)	-0.12	-0.22	0.04	-0.04	0.08	0.09
	[0.27]	[0.26]	[0.25]	[0.33]	[0.17]	[0.15]
Y (lag 4)	-0.19	-0.16	-0.13	-0.14	0.30**	0.32**
	[0.34]	[0.31]	[0.20]	[0.21]	[0.13]	[0.13]
qed (lag 1)		0.53		0.12		-0.10
		[0.34]		[0.51]		[0.31]
qed (lag 2)		-0.50*		0.32		0.36
		[0.28]		[0.54]		[0.29]
qed (lag 3)		0.24		-0.48		0.08
		[0.51]		[0.66]		[0.40]
qed (lag 4)		-0.22		-0.01		-0.28
		[0.37]		[0.48]		[0.33]
constant	1.11	1.45	-1.58	1.39	0.99	0.67
	[0.84]	[1.09]	[5.14]	[6.89]	[0.68]	[0.79]
Observations	32	32	32	32	32	32
R-squared	0.41	0.47	0.78	0.79	0.92	0.93
H_0: all coefficients on X_{t-i} are zero						
F-test statistic p-value		0.33		0.64		0.45

* p<0.10, ** p<0.05, *** p<0.01

Note: Robust standard errors are reported in square brackets.

In this table i=1, 2, 3, 4

Table A7: Assumptions needed for unbiased and consistent estimators in fixed effects regressions

Assumption 1
For each state *i* in each time period *t* the model can be written as, $$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + u_{it}, \qquad where\ u_{it} = a_i + v_{it}$$ where a_i is constant across time but differs between individuals and v_{it} is a random idiosyncratic error.
Assumption 2
Sampling across the cross-sections is random.
Assumption 3
There is variation in the explanatory variables for at least one state *i* and there exists no perfect linear relationship between any of the explanatory variables.
Assumption 4
For each time period t, the expected value of the idiosyncratic error given the explanatory variables in all time periods and the unobserved state effect is zero: $$E(u_{it} \mid x_{1it} + \cdots + x_{kit}, a_i) = 0$$
Under assumptions 1-4, the estimators in the fixed effects regression are unbiased and consistent as the number of periods T remains fixed and the sample N→∞.