Institutions, Education, and Economic Performance

Jonathon Adams-Kane and Jamus Jerome Lim*

September 21, 2008

Abstract

This paper considers the interactions between governance, educational outcomes, and economic performance. More specifically, we seek to establish the linkages by which institutional quality affect growth by considering its mediating impact on education. While the contribution of both human capital and institutions to growth are often acknowledged, the channels by which institutions affect human capital and, in turn, growth, has been relatively underexplored. Our empirical approach adopts a two-stage strategy that estimates national-level educational production functions which include institutional governance as a covariate, and uses these estimates as instruments for human capital in cross-country growth regressions.

KEYWORDS: Institutions, human capital, education, economic growth
JEL CLASSIFICATION: H11, O15, O43

*University of California, Santa Cruz, and the World Bank. Emails: fujax9@ucsc.edu and jlim@worldbank.org. If we had to use one word to describe the preliminary stage that this paper is in, it would be "sushi"; please do not quote or cite. This paper was conceived over many conversations at the chief economist office of the Human Development Network at the Bank. We thank especially Thorsten Janus, Maureen Lewis, and Gunilla Pettersson for early comments. Financial support has thus far come entirely out of our own pockets, but we will willingly receive any that come our way. The findings, interpretations, and conclusions expressed in this article are entirely those of the authors. They do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

1 Introduction

One of the enduring puzzles in the study of human capital and income has been the apparent inconsistency between the empirical micro- and macro-econometric evidence. Studies using Mincer (1974)-style earnings functions generally find that educational levels is one of the strongest predictors of lifetime income, but this intuitive result does not generally survive aggregation: Educational attainment is, by most measures, largely unrelated to national income.

Earlier studies that considered the contribution of human capital to growth (Barro 1991; Mankiw, Romer & Weil 1992) typically found a large and significant influence of such capital—as proxied by enrollment rates—on income per capita. However, later papers (Benhabib & Spiegel 1994; Pritchett 2001) have not only found an insignificant contribution, but in some cases have actually established a *negative* relationship between human capital and income.

This stands in stark contrast to a very large body of microeconometric labor research that has found a strong and persistent relationship between educational levels and wage rates. Although estimates are noisy and may depend on the time period chosen, the general result that earnings increase linearly with schooling completion has been found to hold for both U.S. (Heckman, Lochner & Todd 2006) as well as international (Peracchi 2006) data.

This micro-macro incongruence has led to various efforts aimed at resolving the paradox. One approach argues that human capital is either poorly measured or mismeasured. This approach stresses how existing education stock data may either fail to capture important quality dimensions (Behrman & Birdsall 1983; Hanushek & Kimko 2000), or may suffer from systematic data deficiencies (Cohen & Soto 2007; Doménech & de la Fuente 2006). Accounting for these measurement issues would then resolve the paradox.

Another school of thought has stressed the importance of educational governance failures. Factors such as teacher absenteeism, informal payments, and corruption in schools erode the productivity of the education sector (Reinikka & Svensson 2005; Rogers 2008) and reduce the incentives for human capital accumulation (Gupta, Davoodi & Tiongson 2001). This is an institutional failure, which can subsequently spill over into growth outcomes (Acemoglu, Johnson & Robinson 2005; Galor, Moav & Vollrath 2008). Given the poor institutional environment in which learning occurs, the failure of traditional educational statistics to capture the actual stock of human capital is hardly surprising.

These two resolutions are not unrelated; governance failures often imply poor quality of education. Nonetheless, authors have tended to stress one approach over another.[1]

The major challenge in the empirical study of the role of human capital in growth is centered of the endogeneity of human capital. While there is a strong theoretical basis for how human capital can drive growth in both neoclassical (Lucas 1988) and endogenous (Romer 1990) models, there is also the possibility of reverse causality, possibly through a discount rate channel (Bils & Klenow 2000). This endogeneity suggests that naïve attempts to measure the contribution of human capital will encounter a bias in their estimates.

Our empirical approach adopts a two-stage strategy: First, we estimate national-level educational production functions that include institutional governance and inputs to schooling as covariates. Second, we use these estimates from the first stage as instruments for human capital in cross-country regressions of steady-state income. This method not only provides new cross-country estimates of the impact of governance measures on educational outcomes, but also addresses the endogeneity concerns that arise when using direct measures of education in a regressions of this nature.

Moreover, our use of instrumental variables (IV) allows us to reconcile the two major explanations that have been advanced to resolve the micro-macro human capital puzzle. By including governance measures in the education production function, we directly account for the institutional framework in which human capital accumulation occurs. The methodology also allows us to sidestep the concerns surrounding the mismeasurement of human capital, so long as our instruments are chosen carefully and satisfy the necessary validity conditions.

The paper closest in spirit to our own is that of Hanushek & Kimko (2000), who use a similar two-step estimation procedure but estimate a growth equation in the second stage. Unlike these authors, however, we motivate our model directly from a theoretical augmented Solow growth model, and our empirical strategy does not require us to generate projections of unavailable data in order to obtain a sufficiently-sized sample. In addition, we include governance measures that we regard as both theoretically and empirically important for human capital production. Our approach is also complementary to the work of Glaeser, La Porta, López-de Silanes & Shleifer (2004), who also use a two-stage strategy to argue that human capital, rather than institutions, is a stronger predictor of per capita income. Unlike them, we employ a different choice of instruments, and our substantive concern is driven by a neoclassical growth model, rather than a "fundamental determinants" (Rodrik, Subramanian & Trebbi 2004) approach.

Our main results are supportive of the notion that schooling is central to

[1] Pritchett (2001) further argues that the results could be due to stagnant demand for education labor in developing countries. This explanation is less likely, however, given both international (Berman, Bound & Machin 1998) and plant-level evidence that suggests that the demand for skilled labor is reasonably strong in many developing countries (Fajnzylber & Fernandes 2008; Harrison & Hanson 1999; Pavcnik 2003).

economic growth. Our benchmark specifications find that a 1 percent increase in human capital contributes 3.02–3.33 percent to income per capita, and this contribution outstrips that of physical capital. In our robustness tests, we also show that this result survives the inclusion of additional explanatory variables in the second stage, as well as the use of alternative specifications in the first stage, including specifications allowing governance to be endogenous to income and/or endogenous to human capital.[2] We also demonstrate that the main results follow even when we alter our specification to exploit panel data.

Our findings are of considerable academic and policy interest. Empirical studies of human capital have frequently been hampered by the difficulty of isolating the causal impact of education on per capita income. Furthermore, to the extent that institutions are themselves subject to change, corroborating the body of microeconomic evidence on governance and education provides further impetus for institutional reform in developing countries.

The rest of the paper is organized as follows. Section 2 will present the motivating theoretical model. We then report the empirical results in Section 3, before a final section concludes with policy implications.

2 A Simple Model of Growth, Human Capital, and Governance

Our motivating theoretical model is an augmented Solow (1956) growth model, expanded to allow for three reproducible factors: Labor, L, physical capital, K, and human capital, H (Mankiw *et al.* 1992). Output at time t is generated by the production function

$$Y_t = K_t^{\alpha} H_t^{\beta} \left(A_t L_t\right)^{1-\alpha-\beta}, \quad 0 < \alpha, \beta < 1, \tag{1}$$

where A is the current level of (exogenous) technology, and we assume decreasing returns to all capital, so that $\alpha + \beta < 1$.

The microeconomic literature on the education production function (Todd & Wolpin 2003) argues that cognitive achievement for a given individual i is determined by innate ability, η, family inputs, F, and school inputs, S. At the individual level, human capital at time t is therefore a function

$$H_{it} = h\left(\eta_i, F_{it}, S_{it}; G_t\right),$$

where G is the (exogenous) institutional environment whereby learning takes place, and we assume that individual ability is time-invariant. Aggregating over all effective units of labor gives

$$\begin{aligned} H_t &= \int_1^{A_t L_t} h\left(\eta_i, F_{it}, S_{it}; G_t\right) di \\ &= F_t \gamma S_t^{\epsilon} \left(A_t L_t\right)^{1-\gamma-\epsilon} \cdot G_t^{\phi}, \quad 0 < \gamma, \epsilon < 1, \end{aligned} \tag{2}$$

[2] Lipset (1960) argues that both economic growth and human capital accumulation cause institutional change, a hypothesis supported by Glaeser *et al.* (2004).

where we further assume a Cobb-Douglas form and decreasing returns to inputs with $\gamma + \epsilon < 1$. Note the omission of the ability term at the aggregate level; this amounts to assuming that innate ability is distributed normally across countries at the global level, such that there are no significant cross-country differences. Taking logarithms of (2) gives the (steady-state) amount of human capital per effective unit of labor:

$$\ln\left[\frac{H_t}{L_t}\right] = \ln A_0 + gt + \gamma \ln f + \epsilon \ln s + \phi G, \tag{3}$$

where we follow convention and rewrite $f \equiv \frac{F}{AL}$ and $s \equiv \frac{S}{AL}$ in intensive form, representing family and school inputs per unit of effective labor.

Technology progresses and labor grows at exogenous rates described by

$$A_t = L_0 e^{gt}, \quad L_t = L_0 e^{nt},$$

giving capital accumulation according to the ordinary differential equations

$$\dot{k}_t = s_k y_t - (n + g + \delta)\, k_t, \tag{4a}$$

$$\dot{h}_t = s_h y_t - (n + g + \delta)\, h_t, \tag{4b}$$

where s_k and s_h are, respectively, the investment shares of physical and human capital, δ is the rate of capital depreciation, and as before $y \equiv \frac{Y}{AL}$, $k \equiv \frac{K}{AL}$, and $h \equiv \frac{H}{AL}$ are in intensive form.

The steady state levels of physical and human capital are straightforward, and given by

$$k^* = \left[\frac{s_k^{1-\beta} s_h^{\beta}}{n + g + \delta}\right]^{\frac{1}{1-\alpha-\beta}}, \quad h^* = \left[\frac{s_k^{\alpha} s_h^{1-\alpha}}{n + g + \delta}\right]^{\frac{1}{1-\alpha-\beta}}.$$

Substitution into (1), taking logarithms, and re-substituting the steady-state share of human capital back into the resulting equation yields steady-state income per worker given by

$$\ln\left[\frac{Y}{L}\right] = \ln A_0 + gt + \frac{\alpha}{1-\alpha} \ln s_k + \frac{\beta}{1-\alpha} \ln h^* - \frac{\alpha}{1-\alpha} \ln (n + g + \delta). \tag{5}$$

Together, (3) and (5) are the system of two equations that we take to the data.

3 Empirical Tests of Income, Education, and Institutions

3.1 Empirical Model

Our empirical model is based on the system of equations summarized by (3) and (5):

$$\ln\left[\frac{H_{it}}{L_{it}}\right] = \theta_0 + \mu_i + \theta_1 G_{it} + \ln\left[\frac{\mathbf{F}_{it}}{L_{it}}\right]\mathbf{\Theta}_2 + \ln\left[\frac{\mathbf{S}_{it}}{L_{it}}\right]\mathbf{\Theta}_3 + \varepsilon_{it}, \tag{6}$$

$$\begin{aligned}\ln\left[\frac{Y_{it}}{L_{it}}\right] &= \pi_0 + \rho_i + \pi_1 \ln s_{k,it} + \pi_2 \ln\left[\frac{H_{it}}{L_{it}}\right] - \pi_3 \ln\left(n+g+\delta\right) \\ &\quad + \mathbf{X}_{it}\mathbf{\Pi}_4 + \nu_{it},\end{aligned} \tag{7}$$

where G_{it} is governance, $\mathbf{F}_{it}$ and $\mathbf{S}_{it}$ are vectors of family and school inputs to human capital production for country i at time t, respectively, H_{it} is human capital, $s_{k,it} = \frac{I_{it}}{Y_{it}}$ is the investment share of GDP, $(n+g+\delta) = n+0.05$ is the net rate of depreciation of effective units of labor,[3] $\mathbf{X}_{it}$ is a vector of additional controls, Y_{it} is GDP, μ_i and ρ_i are time-invariant country fixed effects, and $\varepsilon_{it} \sim N\left(0, \sigma_\varepsilon^2\right)$ and $\nu_{it} \sim N\left(0, \sigma_\nu^2\right)$ are i.i.d. disturbance terms. The theoretical prior for our main coefficient of interest, π_2, is positive.

In our robustness section, we populate the vector $\mathbf{X}_i t$ with several other controls that have been found to be important in cross-country growth regressions. Similarly, we have entered family and school inputs as vectors, to accommodate the fact that the education production function literature has identified a host of possible candidates for important inputs to student achievement. In our benchmark specifications, we maintain parsimony with only one input for F and S; we relax this restriction in our robustness section.

3.2 Estimation and Identification Strategy

In our benchmark tests, we employ three main variables in our first-stage regressions. Of these three, two are plausibly exogenous, and could function as instruments; the third may suffer from simultaneity concerns, and is only used in conjunction with our other instruments.

Our first, and primary, instrument is government effectiveness. Although there are potentially many channels by which an effective government bureaucracy can affect economic outcomes, we contend that the primary means by which this occurs is through service delivery, and in particular the delivery of educational services. In many countries, especially developing ones, educational expenditure is one of—if not the—largest components of total public expenditure, and education is largely publicly-provided. If government effectiveness does matter to growth, there is a strong likelihood that it does so mainly through

[3] We follow Mankiw *et al.* (1992) and assume that g and δ are constant across countries and their sum is approximated by calibrated data of 0.02 and 0.03, respectively.

its mediating effect on the delivery of education. We visually capture the relationship between governance and human capital in Figure 1.

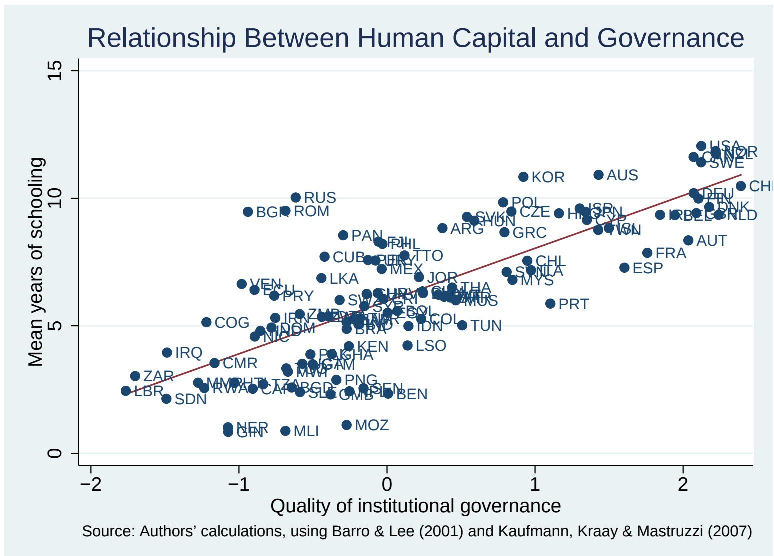

Figure 1: Positive relationship between quality of institutional governance and mean years of schooling, 2000, with fitted regression line. The (bivariate) regression is significant at conventional levels.

The other main channel by which effective government may affect economic outcomes is through policy, especially macroeconomic policy. We are less concerned that this is an issue, however, because there is fairly abundant evidence that policy variables do not exert a systematic influence on economic growth (Levine & Renelt 1992; Sala-i-Martin 1997).[4] In order to rule our any remaining simultaneity concerns, we use the lagged effectiveness variable. Overall, we are reasonably confident that government effectiveness satisfies the exclusion restriction in the first stage. For completeness, however, we provide a formal test of the strength of this particular assumption when we discuss the benchmark results.

The second instrument that we use is the consumption-investment ratio, which acts as a proxy for family inputs into education. To the extent that household educational expenditures is an investment good, the C/I ratio offers a plausibly exogenous instrument for family inputs that is not, theoretically, systematically related to the level of income per capita. While an obvious candidate for household inputs is income per capita, it is essentially the same as

[4]This should perhaps be qualified. There is some evidence that *bad* policy choices—such as financial repression or severe trade restrictions—may negatively affect country performance. However, policies that can be directly associated with government effectiveness—such as monetary and fiscal policy—tend to be insignificant in standard cross-country growth regressions.

the left-hand-side variable in the second stage regression, and thus clearly not exogenous.

Our final variable is the pupil-teacher ratio, which is our proxy for school inputs. We choose this variable, instead of other candidates, in part due to the strong case made for class size as a key determinant of schooling outcomes due to school resources (Krueger 2003), and in part because of its availability across countries and time. There are some legitimate concerns of simultaneity bias in including this variable: Countries with higher incomes per capita are likely to be able to afford to increase schooling resources, lowering the pupil-teacher ratio. Without a measure of school inputs, the tradeoff is efficiency of the estimates due to a poorer fit in the first stage; we report specifications with and without the inclusion of this variable.

The remaining endogeneity issue is that of omitted variable bias. While it is possible that government effectiveness or the consumption-investment ratio can influence income per capita through an intervening omitted variable, or is affected by an omitted variable that also affects income per capita, this is not suggested by our theoretical model. Moreover, we are inclined toward a fairly parsimonious model, given the general lack of robustness of other, atheoretical explanatory variables that have been advanced in the literature. However, we take steps to address this issue in our robustness section.

Estimation of the model is via two-stage least squares, using two-step generalized method of moments (GMM) and adjusted for heteroskedasticity-robust standard errors. For robustness tests using panel data, we run both fixed effects IV-GMM with correction for heteroskedasticity, clustering, and serial correlation, as well as system GMM using the orthogonal deviations transformation for the endogenous regressors (Arellano & Bover 1995) and Windmeijer-corrected standard errors. In most of our specifications, our model is overidentified, and we accordingly report the Hansen J-test of overidentifying restrictions.

3.3 Data Description

Our cross-country macroeconomic data are drawn mainly from the World Bank's *World Development Indicators*. We supplement these with data from several other sources. Our primary measure of the human capital stock is the Barro & Lee (2001) dataset on educational attainment. Our supplementary educational data were mainly from the UNESCO Institute for Statistics' *Global Education Statistics* database. Our primary governance data were the Worldwide Governance Indicators (Kaufmann, Kraay & Mastruzzi 2007), which not only provides disaggregation into the subcomponents that we need, but are also, in our view, the highest-quality data available.

The specific measures employed, as well as other data sources and additional controls used in the robustness tests, are described in full in the data appendix.

3.4 Main Findings

In Table 1 we report the main results of our benchmark model. Specification (*B1*) is the least squares estimates for the augmented Solow model consistent with (7). The sample comprises 103 countries, and the model provides a reasonably good fit. The human capital contribution is statistically significant, and enters with the expected sign. However, endogeneity concerns lead us to discount these results.

Table 1: Benchmark regressions of GDP per capita†

	(*B1*)	(*B2*)	(*B3*)	(*B4*)	(*B5*)	(*B6*)
	Second stage income equation					
Investment share	0.432	0.836	1.097	-0.002	0.689	0.255
	(0.34)	(0.47)*	(0.48)**	(0.42)	(0.27)**	(0.32)
Net rate of depreciation	-0.900	0.815	0.801	1.889	0.744	1.695
	(0.63)	(0.99)	(1.02)	(0.74)**	(0.98)	(0.69)**
Human capital	1.840	3.125	3.142	3.329	3.024	3.250
	(0.23)***	(0.48)***	(0.44)***	(0.39)***	(0.41)***	(0.32)***
Constant	4.111	7.231	7.547	8.616	6.954	8.545
	(1.58)**	(2.35)***	(2.56)***	(1.84)***	(2.33)***	(1.79)***
	First stage human capital equation					
Family resources		-0.359	-0.753		-0.377	
		(0.29)	(0.32)**		(0.24)	
School resources		-0.557	-0.626	-0.548		
		(0.20)***	(0.22)***	(0.17)***		
Governance		0.136		0.137	0.251	0.277
		(0.06)**		(0.05)***	(0.05)***	(0.04)***
Broad governance			0.116			
			(0.07)*			
Constant		1.657	1.749	1.760	-1.971	-1.694
		(1.39)	(1.50)	(1.03)*	(0.96)**	(0.71)**
Adj R^2	0.715	0.534	0.498	0.434	0.591	0.508
Anderson LR		31.544***	29.135***	40.837***	27.779***	39.049***
Cragg-Donald F		12.315	11.252	24.789	15.615	45.639
Hansen J		1.717	1.535	0.255	0.032	-
N	103	64	60	83	78	103

† Notes: Huber-White (robust) standard errors reported in parentheses. First stage regressions included second stage controls as instruments, but are not reported. Hansen statistics for exactly identified models are replaced with a dash. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

The top half of column (*B2*) reports the IV estimates for the baseline specification. In this specification, we use the pupil-teacher ratio as a proxy for school inputs, and the consumption-investment ratio as a proxy for family inputs. Due to data limitations, the full sample falls to 64 countries. Our main coefficient of interest, π_2, remains positive and statistically (and economically) significant. The contribution of capital is also consistent with the theoretical prior, but only marginally significant. The Sargan-Hansen J statistic ($\chi^2 = 2.59, p = 0.27$) indicates that the instruments are valid. The Anderson LR statistic for underidentification is significant, and the Cragg-Donald F for weak instruments is reasonably high ($F = 12.32$, Stock-Yogo $F^{crit} = 9.08$ for 10% relative bias); both suggest that the instruments satisfy the relevance condition. Finally, the partial

R^2 of the first-stage regression (not reported) is reasonably strong ($R^2 = 0.39$); since there is only one endogenous regressor, this result further corroborates the test for weak indentification ($F = 9.78, p = 0.00$).

The bottom half of column (*B2*) reports the corresponding first stage results. While these estimates are of secondary interest, we note that the coefficients are consistent with the expected signs (recall that the pupil-teacher ratio is expected to be negatively related to human capital), and both school inputs and governance are significant at the 5% level. Finally, it is helpful to point out that, unlike Rogers (2008), our empirical strategy introduces the governance dimension directly as a covariate into the education production function, instead of separating the data into subsamples according to their level of governance. Besides being implied by our theoretical model of Section 2, we also regard this approach as a more direct test of the role that institutional governance might (or might not) play in the determination of human capital accumulation.

For reasons of identification, we have chosen to restrict our measure of governance to government effectiveness. Other than econometric reasons, there is a theoretical reason for doing so. The use of the more comprehensive definition of governance runs the risk of being tautological: If good institutions are *defined*, *ex ante*, as those structures and mechanisms that are most likely to enhance growth, then it is small wonder that, *ex post*, institutions are found to directly affect growth. Governance then becomes significant because we have defined it to be so. However, in order to allay concerns regarding the possibility that our choice of governance indicators are *ad hoc*, in column (*B3*) we repeat the above specification, but with one change: We expand the governance measure to all the six dimensions listed in Kaufmann *et al.* (2007). Our results are essentially unchanged. However, the adjusted R^2 for the first stage is lower, and the coefficient in this case is only weakly significant. We consider this a validation of our choice of a narrower definition of governance.

To account for remaining econometric concerns concerning our choice of instruments, we take three further steps: First, we exclude family inputs altogether, treating all measures of income as endogenous to the model. Second, we exclude school inputs, which as we discussed earlier may suffer from simultaneity bias. Third, we exclude all family and school inputs and rely solely on governance to identify the effect of human capital on income level and growth. These are reported in columns (*B4*) through (*B6*), respectively. The coefficient π_2 remains robust through these three changes, although these are not directly comparable due to changes in the sample size that result from differential data availability.

Taken together, the IV results reported in Table 1 suggest that a 1 percent increase in human capital contributes between 3.02–3.33 percent to income per capita. By way of comparison, physical capital—the only other control variable to feature some significant coefficients across the different specifications—has a contribution that is about three to five times smaller, ranging from 0.69–1.10 percent. As is common for cross-country growth regressions, the large and significant constant term suggests that a substantial unexplained component remains.

These specifications also satisfy the primary diagnostic tests for instrument validity. We note that the Hansen J cannot be computed for specification (*B6*), since the specification is just identified; this specification thus relies on the validity of the exclusion restriction (as discussed in Subsection). To formally test the validity of this important assumption, we exploit a recent procedure developed by Kraay (2008), which utilizes Bayesian inference to explicitly characterize the extent to which prior uncertainty about the assumption affects the posterior distribution of π_2.[5]

We report these tests in Table 2, for differing assumptions with regard to the strength of the prior belief that the exclusion restriction holds exactly. This strength is given by the parameter ω, with higher (lower) values representing greater (lesser) certainty that the exclusion restriction is valid. The supports—for the 2.5th and 97.5th percentiles—are chosen to correspond to a 95 percent confidence interval; changes in the interquantile range are also reported.

Table 2: Tests of validity of exclusion restriction for governance[†]

	$\omega = 5$	$\omega = 10$	$\omega = 100$	$\omega = 200$	$\omega = 500$	$\omega = \infty$
	Posterior distribution for π_2					
2.5th percentile	1.49	2.02	2.70	2.74	2.80	2.82
Mode	3.52	3.54	3.53	3.54	3.55	3.55
97.5th percentile	6.12	5.62	4.83	4.86	4.80	4.75
Change in interquantile range	4.63	3.60	2.13	2.12	2.00	1.93

[†] Notes: Posterior distributions calculated assuming that the distribution of prior probabilities that the exclusion restriction holds at 10% level. Corresponding supports are |0.46|, |0.34|, |0.12|, |0.08|, |0.05|, and 0, respectively.

Relative to the case where there is no prior uncertainty about the exclusion restriction ($\omega = \infty$), the supports for the posterior distribution widens (from 1.93 to 4.63) as there is greater uncertainty ($\omega \to 5$), as expected. However, the mode remains stable, and even in the case of extreme uncertainty about the validity of the exclusion restriction ($\omega = 5$), the interval does not include zero, signifying the strength of the instrument. An alternative way of looking at this result is captured in Figure 2; here, while greater uncertainty over instrument validity leads to a wider dispersion in possible π_2 values, this change in the distribution is sufficiently small that the contribution of human capital continues to matter.[6]

[5] The details of the analysis are described briefly in Appendix A.2.

[6] An important consideration of the tests are what the results would be *if* the distribution of priors was not centered on zero; in particular, if it were centered on a positive value. In this case, Kraay (2008) suggests that the nonzero mean would need to be subtracted out from the posterior distribution, which would result in a lower value for the 2.5th percentile that may include zero. However, since we do not have a means of reliably estimating this prior, we can only allude to this possibility as an important caveat to the results above.

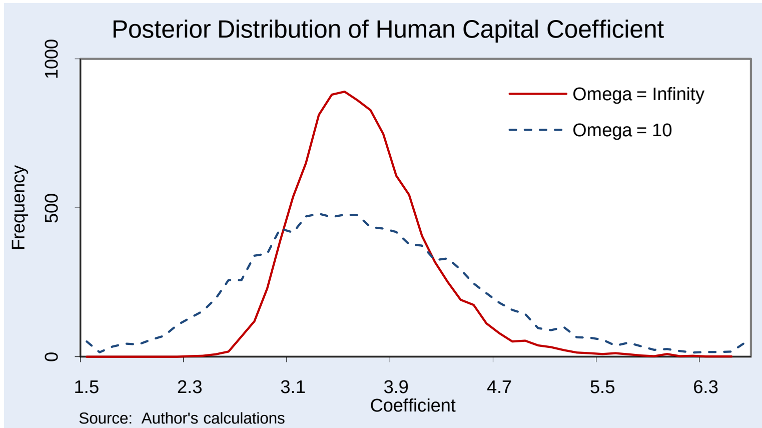

Figure 2: Posterior distribution for coefficient of human capital, with alternative assumptions about the validity of the exclusion restriction. Lower values of ω indicate greater prior uncertainty that the instrument satisfies the orthogonality condition. Even with high levels of uncertainty, the posterior distribution of the slope coefficient does not include zero.

3.5 Robustness Tests

In the benchmark models, we did not introduce any additional controls to explain cross-country income per capita. Here, we allow $\mathbf{X}$ to include variables that the literature has identified as important. More specifically, we draw on a selection of the variables that Levine & Renelt (1992) and Sala-i-Martin (1997) argue are robust empirical relations: The trade share of GDP, geographic location, and infrastructure.[7] To this we include some relatively more recent candidates in the empirical growth literature: Ethnolinguistic fractionalization (Easterly & Levine 1997), democratic development (Barro 1996), and social capital (Knack & Keefer 1997). These are reported in columns (*R1*)–(*R6*) of Table 3.

The significance of the coefficient on human capital survives the inclusion of all these additional controls. As before, while the coefficients are not directly comparable, we note that the human capital contribution is statistically and economically significant, with a range $[2.65, 4.08]$. The coefficient on physical capital is occasionally statistically significant, but its contribution is never greater than 1.53 percent, and is always dominated by the human capital contribution. None of the other variables that have been identified as important enter significantly. Also, the instruments pass both the under- and over-identification tests, and in most cases satisfy the tests for weak instruments as well.

[7] We used road density as a proxy for infrastructure, but we also explored alternative proxies such as the share of rural population and a weighted average of the percentage of population with access to water and sanitation facilities. Our qualitative results were affected by these alternatives.

Table 3: Regressions of GDP per capita with additional controls†

	(R1)	*(R2)*	*(R3)*	*(R4)*	*(R5)*	*(R6)*
Investment share	0.931 (0.51)*	0.752 (0.44)*	0.663 (0.74)	0.879 (0.47)*	1.527 (0.73)**	1.195 (0.47)**
Net rate of depreciation	0.853 (1.00)	0.997 (0.88)	0.779 (1.13)	0.935 (1.10)	1.303 (1.56)	0.444 (0.89)
Human capital	3.160 (0.49)***	2.901 (0.49)***	2.992 (0.54)***	3.203 (0.50)***	4.076 (0.77)***	2.646 (0.44)***
Trade share	-0.092 (0.14)					
Geography		0.152 (0.10)				
Infrastructure			0.079 (0.10)			
Ethnolinguistic fractionalization				0.131 (0.20)		
Social capital					1.541 (1.51)	
Democracy						0.206 (0.14)
Constant	7.805 (2.56)***	7.565 (2.11)***	6.854 (3.13)**	7.400 (2.74)***	6.924 (3.78)*	7.326 (2.33)***
Adj R^2	0.517	0.590	0.498	0.478	0.523	0.678
Anderson LR	31.099***	28.409***	20.051***	24.981***	17.367***	34.238***
Cragg-Donald F	11.888	10.636	7.044	9.084	5.984	13.677
Hansen J	1.702	2.148	1.957	1.320	0.980	2.773
N	64	63	54	63	39	58

† Notes: Huber-White (robust) standard errors reported in parentheses. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

We now proceed to consider alternative variables for and permutations of our exogenous instruments.

An alternative way to qualify family inputs in the education production function is to recognize that families with a greater share of parental authority invested in the mother—usually due to higher levels of education attained by them—are more likely to invest a greater share of family resources on education (Carneiro, Meghir & Parey 2007). We use this variable as an additional instrument to proxy for family inputs. We report this specification in column (*Z1*) of Table 4. In this case, the instruments are somewhat weak, but human capital remains positive and significant.[8]

Some authors have recently made a case for how genetic factors may influence growth, either in terms of genetic diversity (Ashraf & Galor 2008) or, more specifically, through the general intelligence quotient factor g (as either as a proxy for human capital (Jones & Schneider 2006) or as an indicator of unobservable individual ability in the process of human capital formation (Weede & Kämpf 2002)). There have been numerous criticisms of the use of g as a reliable indicator of general intelligence.[9] For our purposes, it is sufficient to note two

[8] We also explored *replacing* the family input variable altogether, and while our qualitative results were unchanged, the instrument set did not satisfy the exclusion condition.

[9] We will not delve too deeply into the large (and contentious) literature on the psychometric measurement of intelligence and cognitive ability. Devlin, Fienberg, Resnick & Roeder (1997)

Table 4: Regressions of GDP per capita with alternative controls†

	(*Z1*)	(*Z2*)	(*Z3*)	(*Z4*)	(*Z5*)
Investment	1.636	0.550	0.013	-1.067	0.970
share	(0.49)***	(0.23)**	(0.51)	(0.10)***	(0.42)**
Net rate of	1.672	0.821	4.464	-1.313	0.306
depreciation	(1.17)***	(0.65)	(1.07)***	(0.61)**	(0.82)
Human	3.556	3.242		1.788	3.073
capital	(0.48)***	(0.36)***		(0.20)***	(0.40)***
Alternative			7.981		
human capital			(0.85)***		
Constant	10.827	6.490	-17.418	0.491	6.067
	(1.75)***	(1.72)***	(3.07)***	(1.46)	(1.92)***
Adj R^2	0.528	0.649	0.173	0.832	0.541
Anderson LR	26.546***	53.722***	31.075***	36.005***	47.987***
Cragg-Donald F	7.460	39.036	18.248	15.237	15.911
Hansen J	4.955***	2.776	0.276	4.775	12.681***
N	54	63	68	11	64
	(*Z6*)	(*Z7*)	(*Z8*)	(*Z9*)	(*Z10*
Investment	0.761	0.972	0.895	0.260	0.805
share	(0.43)*	(0.37)***	(0.41)**	(0.76)	(0.49)*
Net rate of	0.853	0.997	0.779	0.935	1.303
depreciation	(1.00)	(0.88)	(1.13)	(1.10)	(1.56)
Human	3.070	3.139	2.538	1.173	1.897
capital	(0.46)***	(0.48)***	(0.75)***	(1.80)	(1.14)*
Governance			0.203	0.602	0.320
			(0.22)	(1.65)	(0.55)
Constant	6.886	7.493	7.096	4.182	5.700
	(2.20)***	(2.29)***	(1.98)***	(1.93)**	(1.79)***
Adj R^2	0.550	0.529	0.696	0.835	0.812
Anderson LR	31.628***	31.680***	9.507***	0.541	1.091
Cragg-Donald F	9.108	9.127	3.096	0.150	0.245
Hansen J	1.907	1.933	1.521	3.765*	4.395
N	64	64	64	34	64

† Notes: Huber-White (robust) standard errors reported in parentheses. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

important reservations, both of which we regard as critical.

The first is methodological. The theoretical foundation for g is premised on the emergence of a single general factor from hierarchical factor analysis of test scores. The problem with inferring that general intelligence exists as a consequence is that a general factor will always result whenever the correlation structure of all intelligence tests are positive (Thomson 1916), which is always true by design. The low power of such tests, especially with limited sample sizes, casts doubt as to whether g does truly exist, or even if it does, whether it can be accurately measured with IQ tests.

The second concern is that measures of g and their growth rates are not stable across time; in particular, they demonstrate a positive time trend. These have been extensively documented both between ethnic groups within countries, as well as between countries (Flynn 2007). Although many resolutions have been

provides a good summary of the key issues in the debate.

proposed to explain this effect, persuasive arguments have been advanced that changes in the cognitive or nutritional environment are responsible. Importantly for our purposes, this implies that IQ itself may be endogenous to the level of economic development of a country.

With these reservations in mind, we nonetheless include in our empirical tests a measure of intelligence, due to Lynn & Vanhanen (2002),[10], as a strong proxy for all resource inputs (so that the instrument set includes only IQ and government effectiveness).[11] This is reported in column (*Z2*). As before, our results are largely unchanged.

In the specifications listed in Table 1, we shied away from using attainment data (in the form of test scores). By and large, the international comparability across different test types and time periods are suspect, and where comparable data are available, they are often only for a very limited set of (mostly developed) countries. Moreover, our instrumental variables strategy already accounts for issues of mismeasurement, conditional on our instruments satisfying the necessary exclusion conditions. Nonetheless, we use a recently-compiled database of comparable attainment data (Altinok & Murseli 2007) to examine how our results change when we utilize a more accurate measure of human capital quality. The results are reported in column (*Z3*).[12] Human capital remains significant, and in this case its contribution more than doubles, so that a 1 percent increase in human capital leads to an almost 8 percent increase in output per worker. We do note the far poorer fit of the specification, however, which we feel justifies our decision not to use this measure as our primary measure of human capital.

The microeconometric literature on education production functions suggests that, in addition to the pupil-teacher ratio, several other inputs have been important (Hanushek 2003; Pritchett & Filmer 1999). We include, as additional instruments, a selection of the determinants that have been found to be more consistently significant: The percentage of trained teachers (as a macroeconomic proxy for teacher ability, usually measured with teachers' years of schooling or experience), and public education expenditures (a macroeconomic proxy for resources devoted to teacher salaries and school infrastructure). This specification is reported in column (*Z4*).[13] Although the results are once again similar, we note that the specification suffers from a small sample problem, which may limit inference.[14]

[10]The measures themselves have also been subject to dispute. The source data used in the construction of the dataset have been criticized as being based on excessively small, unrepresentative samples of national populations, and concerns have been raised about the accuracy of the reported scores and about the normalization methods employed to render the scores internationally comparable.

[11]Alternatively, we could have included it in (6) as a measure of innate ability, η, which we now allow to differ between nations. Doing so did not affect the qualitative nature of our results, but the instrument set fails the Hansen J test.

[12]We are again forced, by virtue of satisfying the overidentification test, to exclude family inputs from the instrument set.

[13]The microeconomic literature also finds that teacher quality is a very important source of variation in student performance (Hanushek 2003). Unfortunately, there is close to no international data available for teacher quality.

[14]Other permutations and combinations of these additional school inputs yielded similar

The next three columns, (*Z5*)–(*Z7*), introduce interaction terms between governance and resource inputs. These are for governance and school inputs, governance and family inputs, and family and school inputs, respectively. Although not fully justified by our theoretical model, the interaction term allows for the possibility that the efficacy of school inputs may be conditional on the institutional environment. This is intuitively plausible, and the interaction term also serves as a possible instrument that is orthogonal to the error term in the second stage. Adding these interaction terms, however, does not modify our principal conclusions concerning the coefficient for human capital, which remains relatively stable throughout. Note, however, that (*Z5*) does not satisfy the overidentification test.

Our final three specifications endogenize potentially the most problematic instrumental variable: Government effectiveness. Column (*Z8*) uses lagged government effectiveness (from 1996) as an instrument for contemporaneous (year 2000) governance. The magnitude of the human capital contribution falls, but remains significant at the 1 percent level, while the coefficient for physical capital is also significant at the 5 percent level. Interestingly, government effectiveness is *insignificant* when included in the second stage, while lagged effectiveness is significant and positively signed in the first stage human capital equation. This gives us confidence that the effects of good governance—at least when measured with government effectiveness—operates primarily through its mediating role on human capital. This is also the argument first raised in Glaeser *et al.* (2004), although they arrive at their claim from a different angle. It is also consistent with the work of Galor *et al.* (2008), who argue that the Great Divergence can be attributed, in part, to the emergence of institutions that promote the formation of human capital.

The fairly large literature that has emerged following Acemoglu, Johnson & Robinson (2001) has utilized, as instruments for institutions, settler mortality. We are somewhat reluctant to use these instruments, however, for two reasons. First, while a convincing case can be made for how the historical disease environment is a plausibly exogenous instrument for contemporary property rights institutions—or broader definitions of institutions—the linkage is, in our view, weaker when institutions are defined, as we do here, as the efficacy of the current government bureaucracy. Second, recent work has questioned the quality of the settler mortality data (Albouy 2008), and corrections to these data leads to settler mortality becoming a weaker instrument.

For consistency with the rest of the literature, however, we include in our instrument set settler mortality (column (*Z9*)) and, following Hall & Jones (1999), instruments corresponding to the fraction of the population of European descent column (*Z10*) (we maintain as instruments family and school inputs). As expected, the quality of the combined instrument set is suspect: Specification (*Z9*) fails *both* the exclusion and relevance conditions, and specification (*Z10*) does not pass the underidentification test. In addition, the Cragg-Donald F

significant coefficients for human capital, but typically did not satisfy the overidentification test.

statistics suggest that the instruments are extremely weak. Human capital does show up marginally significant in the latter specification, however, and governance remains an insignificant predictor of income.

3.6 Panel Results

Due to data limitations, the estimates that have been presented thus far have been cross-sectional in nature. It is possible to expand the sample to a panel, but it is important to keep in mind two considerations. First, while the educational attainment data are available for five-year intervals from 1960–2000, the panel is unbalanced, and consequently the 116-country sample has an average of only 4 observations per country. We report the fixed effects regression, analogous to (*B1*), in column (*P1*) of Table 5.[15]

Table 5: Panel regressions of GDP per capita†

	(*P1*)	(*P2*)	(*P3*)	(*P4*)	(*P5*)	(*P6*)
			Second stage income equation			
Investment	0.162	0.152	0.111	0.126	0.031	0.139
share	(0.05)***	(0.04)***	(0.08)	(0.08)	(0.11)	(0.10)
Net rate of	-0.098	0.043	0.337	0.349	0.733	-0.612
depreciation	(0.13)	(0.13)	(0.18)*	(0.19)*	(0.28)***	(0.46)
Human	0.409					
capital	(0.11)***					
Alternative		0.323	1.503	1.546	2.183	-0.937
human capital		(0.05)***	(0.44)***	(0.47)***	(0.72)***	(1.12)
Constant	8.263	8.101				
	(0.39)***	(0.37)***				
			First stage human capital equation			
Family			0.029	0.309		0.011
resources			(0.11)	(0.11)		(0.09)
School			-0.253	-0.261	-0.180	
resources			(0.08)***	(0.08)***	(0.06)***	
Governance			-0.047		-0.030	-0.081
			(0.06)		(0.04)	(0.06)
Broad				-0.032		
governance				(0.07)		
F	9.018***	21.990***	6.173***	5.980***	5.478***	1.261
Anderson LR			13.012***	12.015***	12.627***	4.395*
Cragg-Donald F			4.356	4.017	6.342	2.188
Hansen J			3.256	4.024	2.407	1.294
N	511	1175	435	435	658	536

† Notes: Heteroskedasticity, cluster, and autocorrelation-robust (asymptotic) standard errors reported in parentheses. With the exception of the pooled specification, regressions included country and time fixed effects. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

Second, given that the governance and educational attainment data overlap for only one year (2000), we need to use an alternative measure of human capital if we wish to expand the panel in a way that allows us to preserve the use of government effectiveness as an instrument. We do so by substituting our human

[15] The Hausman test detects systematic differences between coefficients and hence a preference for fixed over random effects.

capital measure with data on enrollment rates. The panel with enrollment rates alone is much larger—176 countries, with an average of 7 years—and for reasons of comparability we report the fixed effects regression using this human capital measure in column (*P2*).

The coefficients for human capital in both of these specifications are relatively small: 0.409 and 0.323, respectively, although both are statistically and economically significant. Physical capital also appears significant in both of these specifications, although the magnitudes of their coefficients are also correspondingly smaller. As before, however, we discount these estimates because of endogeneity concerns.

Our benchmark panel, which uses enrollment data but is otherwise analogous to (*B2*), is reported in column (*P3*). It comprises 95 countries, with an average of about 5 time periods per country. As noted in the introduction, the danger that enrollment is a poor proxy measure for human capital is less of a concern as long as our instruments are valid. The Anderson and Hansen tests confirm that this is indeed the case, although it is important to point out that we are forced to use contemporaneous (instead of lagged) government effectiveness as an instrument; it is perhaps for this reason that in the coefficient on governance in the first stage is indistinguishable from zero.

The results largely corroborate the findings of the cross section estimates, with the coefficients on human capital being statistically significant. While the magnitude of the contribution is somewhat smaller, it is still economically significant: A 1 percent increase in human capital leads to a 1.5 percent increase in per capita income. This decline is probably due to the inclusion of country fixed effects, which would capture a good deal of idiosyncratic country-specific variation.

In columns (*P4*)–(*P6*), we make several minor perturbations to this benchmark. Specification (*P4*) replaces government effectiveness with the broad measure of governance, while columns (*P5*) and (*P6*) limit the instrument set by dropping, respectively, family and school inputs as instruments. While dropping family inputs as an instrument or using the broad measure of governance does not affect our results in any qualitative fashion, the instrument set is weakened considerably by the absence of school inputs. Specification (*P6*) satisfies the relevance condition only marginally, and the utility of the model—as given by the F test—is very low. While we report the estimates in this final model for completeness, we are inclined to heavily discount them in our analysis.

Our final robustness check seeks to endogenize as many of the instruments that we have used as possible; of particular concern is the possibility that governance may be endogenous to the income equation. To do so, we exploit the temporal nature of the panel to retrieve internal instruments based on the lags of the endogenous variables. Table 6 reports these results using the panel with enrollment rates as a proxy for human capital, and contemporaneous government effectiveness as the measure of governance.

The specifications are as follows: (*S1*) System GMM estimates of (5), with governance, with one-period lagged GMM-style internal instruments and family and school resources treated as fully exogenous IV-style instruments;[16] (*S2*) Specification (*S1*), but without family and school inputs as exogenous instruments; (*S3*) Specification (*S1*), but with year dummies as additional exogenous instruments; (*S4*) Specification (*S1*), but with two-period lagged GMM-style internal instruments; (*S5*) Specification (*S1*), but with a broad governance measure; and (*S6*) All variables in (3) included as explanatory variables in (5), with one-period lagged GMM-style internal instruments.

Table 6: Regressions of GDP per capita with internal instruments†

	(*S1*)	(*S2*)	(*S3*)	(*S4*)	(*S5*)	(*S6*)
Investment share	-0.317	-0.352	0.189	0.147	0.553	-0.791
	(0.37)	(0.42)	(0.20)	(0.20)	(0.26)**	(0.35)**
Net rate of depreciation	0.190	0.554	0.412	-0.332	-0.043	-0.013
	(1.15)	(0.78)	(1.22)	(0.73)	(0.75)	(0.78)
Human capital	1.651	1.674	1.471	1.277	1.323	0.816
	(0.34)***	(0.32)***	(0.30)***	(0.22)***	(0.18)***	(0.48)*
Governance	0.774	0.519	0.698	0.700		0.352
	(0.26)***	(0.15)***	(0.21)***	(0.13)***		(0.18)*
Broad governance					0.740	
					(0.11)***	
Family resources						-0.550
						(0.29)*
School resources						-1.086
						(0.56)*
Constant	2.255	3.134	4.246	2.984	4.130	8.718
	(3.42)	(2.38)	(2.83)	(1.66)*	(1.73)**	(5.12)*
F	22.405***	30.411***	61.002***	77.530***	48.714***	60.737***
Arellano AR(1)	1.022	0.198	-1.170	-1.381	-0.106	-0.857
Arellano AR(2)	1.016	1.072	0.560	1.377	1.959***	-0.052
Hansen J	39.123	53.451***	45.149	88.030	82.110	63.305
N	445	808	445	445	511	445

† Notes: Heteroskedasticity, cluster, and autocorrelation-robust (asymptotic) standard errors reported in parentheses. A constant term and time dummies were included in the regressions, but not reported. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

We make three comments about the results. First, the instrument set is reasonably sound. With the exception of specification (*S2*), the instruments satisfy the overidentifying restrictions, and the Arellano-Bond test for both AR(1) and AR(2) autocorrelation is satisfied (exempting AR(2) serial correlation in specification (*S5*)). Although not reported, the difference-in-Sargan tests for the (strict) exogeneity of the instrument subsets are generally satisfied.

Second, the coefficient on human capital is significant across all the specifications, ranging from 0.816–1.651 (with the lower bound only marginally significant). As before, the human capital contribution swamps the physical capital share, and in many cases by an order of magnitude.[17] Once again, we have

[16] Strictly speaking, system GMM also uses first differences of endogenous regressors as additional instruments, but this difference structure does not vary since additional lagged differences would lead to redundant moment conditions.

[17] Although investment share is incorrectly signed in some specifications, these estimates are generally statistically indistinguishable from zero. In the one specification where the coefficient on physical capital is significant, it is correctly signed.

validation that human capital is an economically crucial determinant of income patterns.

Third, our measure of governance enters significantly across the different specifications as well, with magnitudes that are about half that of the coefficient on human capital. This stands in contrast to our findings reported in the cross-section (Table 4), and deserves some explanation. The crucial difference to note is that our measure of governance in this case is contemporaneous, rather than lagged, government effectiveness.

Why might this lead to problems? Our estimation method (system GMM) uses weak exogeneity—the assumption that current explanatory variables are not affected by future innovations in income—as an identification strategy. While this may be plausible for human and physical capital, the fact that the current stock of human capital is likely to be affected by past realizations of governance quality means that the simultaneity problem is not completely eliminated when we include current levels of governance as a covariate on the right hand side. In other words, we cannot rule out the possibility that anticipated future levels of income may affect current governance levels, which violates the assumption of weak exogeneity. This may account for the significance of the governance variable, although we cannot completely rule out the possibility that our theoretical model suffers from misspecification concerns.

3.7 Subsample Analysis

Given the centrality of institutional differences, we perform one final set of analyses to tease out the mechanism driving our results. We dissect the panel into subsamples corresponding to the following: (a) The top and bottom fiftieth percentiles of the distribution; (b) Half a standard deviation above and below the mean; (c) One standard deviation above and below the mean, all with respect to the broad institutional governance measure.[18] These are reported in Table 7.

We offer three remarks about the results. First, compromising the sample size typically reduces the strength of the instruments, as reflected in both the over and under-identification tests (especially for the specification in column four), as well as the Cragg-Donald weak instrument tests. This gives us less confidence that endogeneity problems have been fully addressed, and this may also account for the generally smaller point estimates for the coefficient on human capital.

Second, this reduction in sample size also significantly reduces the explanatory power of the model in general. The F statistics in the final two columns are insignificant, as are all the coefficients on the covariates. First stage results (not reported) further suggest a very poor fit for instruments, with low F statistics and insignificant controls.

[18]We also explored subsamples pivoted about the mean, and with larger deviations from the mean, but these subsamples did not yield any additional qualitative insight, and in some cases were not estimable due to small sample sizes.

Table 7: Panel regressions of GDP per capita, by institutional quality†

	$< p50$	$> p50$	$< \mu - \frac{1}{2}\sigma$	$> \mu + \frac{1}{2}\sigma$	$< \mu - \sigma$	$> \mu + \sigma$
Investment	0.115	0.047	0.169	0.152	-0.080	0.139
share	(0.07)*	(0.12)	(0.11)	(0.09)*	(0.16)	(0.11)
Net rate of	0.228	0.155	0.162	-0.148	-0.789	-0.050
depreciation	(0.20)	(0.22)	(0.15)	(0.13)	(1.59)	(0.15)
Human	0.674	2.497	0.568	1.290	0.170	-0.512
capital	(0.29)**	(0.95)***	(0.27)**	(0.78)*	(0.14)	(1.02)
F	5.298***	4.919***	6.069***	5.044***	0.619	0.486
Anderson LR	10.491***	11.354***	17.929***	4.524	16.056***	3.508
Cragg-Donald F	3.501	3.798	6.221	1.471	6.059	1.116
Hansen J	1.882	2.216	1.369	6.780**	0.325	2.085
N	203	224	130	144	30	95

† Notes: Heteroskedasticity, cluster, and autocorrelation-robust (asymptotic) standard errors reported in parentheses. Sample sizes above and below the median differ because not all controls were available for full-sample estimation. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

Third, and most interestingly, human capital appears to matter in institutional environments that are either relatively strong or relatively weak. While this may simply be a consequence of the restricted sample, there is reason to believe otherwise. Subsample regressions that dissect the data into regions or income groups (reported in Appendix A.3) find significant coefficients on human and physical capital, despite some of these subsamples possessing even smaller sample sizes. What is more likely is that countries that fall in the extremes of the institutional quality distribution face systematically different challenges in translating human capital investments into growth outcomes.

For countries with extremely poor quality of institutions—such as Guinea, Laos, and Sudan—improvements in human capital alone are unlikely to make a dent in growth, unless accompanied by institutional improvements that render such investments productive in the context of the broader economy. At the other end of the spectrum, countries that have already accumulated a large stock of human capital—such as Belgium, Finland, and Sweden—may face strong diminishing returns to additional investments in education. While schooling may still matter for lifetime incomes at the individual level, the marginal returns to an additional unit human capital at the country level would be much smaller.

More generally, the results in Table 7 can be interpreted in the light of equations (3) and (5). In countries with low quality of institutions and ineffectual governments (low G), the marginal productivity of effective human capital (h) is likely to low, such that the binding constraint to per capita income growth is in (3). As countries improve their governance levels, this constraint is relaxed, such that human capital makes a positive and significant contribution to income per capita. Finally, for countries with strong institutional frameworks (high G), (3) no longer acts as a constraint to growth. Instead, continued output growth bumps into diminishing marginal productivity, as embodied in the coefficient of human capital in (5).

4 Conclusion

In this paper, we take an alternative approach to reconciling the apparent paradox between micro- and macro-level studies of the role of human capital in income. Specifically, we have argued that the quality of institutions is central to learning and education, so that the role of governance in a country's growth process operates primary though its intervening effect on human capital. Using a range of empirical identification strategies, we have taken this theory to the data, and found support for this conjecture at both the cross-sectional and panel level.

Future research will consider more carefully the mechanisms underlying changes in institutional quality, and its interactions with growth. In particular, by allowing for a dynamic process of institutional change, it may be possible to obtain steady-state expressions for not just human and physical capital, but also institutions, and the interactions between these economic and political factors. Empirical opportunities include directly testing the role of governance in education using micro-level indicators of governance, such as teacher absenteeism rates or the pervasiveness of informal payments in schooling.

References

Acemoglu, K. Daron, Simon Johnson & James A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation". *American Economic Review* 91(5) (December): 1369–1401

Acemoglu, K. Daron, Simon Johnson & James A. Robinson (2005). "Institutions as a Fundamental Cause of Long-Run Economic Growth". In Philippe Aghion & Steven N. Durlauf (editors), *Handbook of Economic Growth*, volume 1 of *Handbooks in Economics*, chapter 6, pp. 385–472. Amsterdam, The Netherlands: Elsevier

Albouy, David Y. (2008). "The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data". *Working Paper W14130*, National Bureau of Economic Research, Cambridge, MA

Alesina, Alberto F., William R. Easterly, Arnaud Devleeschauwer, Sergio Kurlat & Romain T. Wacziarg (2003). "Fractionalization". *Journal of Economic Growth* 8(2) (June): 155–194

Altinok, Nadir & Hatidje Murseli (2007). "International Database on Human Capital Quality". *Economics Letters* 96(2) (August): 237–244

Arellano, Manuel & Olympia Bover (1995). "Another Look at the Instrumental Variable Estimation of Error-Components Models". *Journal of Econometrics* 68(1) (July): 29–51

Ashraf, Quamrul & Oded Galor (2008). "Human Genetic Diversity and Comparative Economic Development". *Discussion Paper Dp6824*, Centre for Economic Policy Research, London, England

Barro, Robert J. (1991). "Economic Growth in a Cross Section of Countries". *Quarterly Journal of Economics* 106(2) (May): 407–443

Barro, Robert J. (1996). "Democracy and Growth". *Journal of Economic Growth* 1(1) (March): 1–27

Barro, Robert J. & Jong-Wha Lee (2001). "International Data on Educational Attainment: Updates and Implications". *Oxford Economic Papers* 53(3) (July): 541–563

Behrman, Jere R. & Nancy Birdsall (1983). "The Quality of Schooling: Quantity Alone is Misleading". *American Economic Review* 73(5) (December): 928–946

Benhabib, Jess & Mark M. Spiegel (1994). "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data". *Journal of Monetary Economics* 34(2) (October): 143–173

Berman, Eli, John Bound & Stephen J. Machin (1998). "Implications Of Skill-Biased Technological Change: International Evidence". *Quarterly Journal of Economics* 113(4) (November): 1245–1279

Bils, Mark & Peter J. Klenow (2000). "Does Schooling Cause Growth?" *American Economic Review* 90(5) (December): 1160–1183

Carneiro, Pedro, Costas Meghir & Matthias Parey (2007). "Maternal Education, Home Environments and the Development of Children and Adolescents". *Working Paper DP6505*, Centre for Economic Policy Research, London, England

Cohen, Daniel & Marcelo Soto (2007). "Growth and Human Capital: Good Data, Good Results". *Journal of Economic Growth* 12(1) (March): 51–76

Devlin, Bernie, Stephen E. Fienberg, Daniel P. Resnick & Kathryn Roeder (editors) (1997). *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve*. Berlin, Germany: Springer

Doménech, Rafael & Angel de la Fuente (2006). "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?" *Journal of the European Economic Association* 4(1) (March): 1–36

Easterly, William R. & Ross Levine (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions". *Quarterly Journal of Economics* 112(4) (November): 1203–1250

Fajnzylber, Pablo & Ana M. Fernandes (2008). "International Economic Activities and Skilled Labor Demand: Evidence from Brazil and China". *Applied Economics* 40(1) (January): 1–15

Flynn, James R. (2007). *What is Intelligence? Beyond the Flynn Effect*. Cambridge University Press

Galor, Oded, Omer Moav & Dietrich Vollrath (2008). "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence". *Review of Economic Studies* (forthcoming)

Glaeser, Edward L., Rafael La Porta, Florencio López-de Silanes & Andrei Shleifer (2004). "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3) (September): 271–303

Gupta, Sanjeev, Hamid R. Davoodi & Erwin R. Tiongson (2001). "Corruption and the Provision of Health Care and Education Services". In Arvind K. Jain (editor), *The Political Economy of Corruption*, chapter 6, pp. 111–141. London, England: Routledge

Hall, Robert E. & Charles I. Jones (1999). "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114(1) (February): 83–116

Hanushek, Eric A. (2003). "The Failure of Input-Based Schooling Policies". *Economic Journal* 112(485) (February): F64–F98

Hanushek, Eric A. & Dennis D. Kimko (2000). "Schooling, Labor-Force Quality, and the Growth of Nations". *American Economic Review* 90(5) (December): 1184–1208

Harrison, Ann E. & Gordon Hanson (1999). "Who Gains from Trade Reform? Some Remaining Puzzles". *Journal of Development Economics* 59(1) (June): 125–154

Heckman, James J., Lance J. Lochner & Petra E. Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond". In Eric A. Hanushek & Finis Welch (editors), *Handbook of the Economics of Education*, volume 1, chapter 7, pp. 307–458. Amsterdam, The Netherlands: Elsevier

Jones, Garett & W. Joel Schneider (2006). "Intelligence, Human Capital, and Economic Growth, A Bayesian Averaging of Classical Estimates (BACE) Approach". *Journal of Economic Growth* 11(1) (March): 71–93

Kaufmann, Daniel, Aart C. Kraay & Massimo Mastruzzi (2007). *Governance Matters: Governance Indicators for 1996–2006*. The World Bank, Washington, DC, vi edition

Knack, Stephen & Philip Keefer (1997). "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation". *Quarterly Journal of Economics* 112(4) (November): 1251–1288

Kraay, Aart C. (2008). "Instrumental Variables Regressions with Honestly Uncertain Exclusion Restrictions". *Policy Research Working Paper 4632*, The World Bank, Washington, DC

Krueger, Alan B. (2003). "Economic Considerations and Class Size". *Economic Journal* 113(485) (February): F34–F63

Levine, Ross & David Renelt (1992). "A Sensitivity Analysis of Cross-Country Growth Regressions". *American Economic Review* 82(4) (September): 942–63

Lipset, Seymour M (1960). *Political Man: The Social Basis of Modern Politics*. New York, NY: Doubleday

Lucas, Robert E., Jr. (1988). "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22(1) (July): 3–42

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

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

Marshall, Monty G. & Keith Jaggers (2005). *Polity IV Project: Political Regime Characteristics and Transitions, 1800–2004*. Center for International Development and Conflict Management, College Park, MD, iv edition

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

Pavcnik, Nina (2003). "What Explains Skill Upgrading in Less Developed Countries?" *Journal of Development Economics* 71(2) (August): 311–328

Peracchi, Franco (2006). "Educational Wage Premia and the Distribution of Earnings: An International Perspective". In Eric A. Hanushek & Finis Welch (editors), *Handbook of the Economics of Education*, volume 1, chapter 5, pp. 189–254. Amsterdam, The Netherlands: Elsevier

Pritchett, Lant H. (2001). "Where Has All the Education Gone?" *World Bank Economic Review* 15(3) (September): 367–391

Pritchett, Lant H. & Deon P. Filmer (1999). "What Education Production Functions Really Show: A Positive Theory of Education Expenditures". *Economics of Education Review* 18(2) (April): 223–239

Reinikka, Ritva & Jakob Svensson (2005). "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda". *Journal of the European Economic Association* 3(2–3) (April/May): 259–267

Rodrik, Dani, Arvind Subramanian & Francesco Trebbi (2004). "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development". *Journal of Economic Growth* 9(2) (June): 131–165

Rogers, Mark Llewellyn (2008). "Directly Unproductive Schooling: How Country Characteristics Affect the Impact of Schooling on Growth". *European Economic Review* 52(2) (February): 356–385

Romer, Paul M. (1990). "Endogenous Technological Change". *Journal of Political Economy* 98(5) (October): S71–S102

Sala-i-Martin, Xavier X. (1997). "I Just Ran Two Million Regressions". *American Economic Review* 87(2) (May): 178–183

Solow, Robert M. (1956). "A Contribution to the Theory of Economic Growth". *Quarterly Journal of Economics* 70(1) (February): 65–94

Thomson, Godfrey H. (1916). "A Hierarchy without a General Factor". *British Journal of Psychology* 8(2): 271–281

Todd, Petra E. & Kenneth I. Wolpin (2003). "On the Specification and Estimation of the Production Function for Cognitive Achievement". *Economic Journal* 113(485) (February): F3–F33

Weede, Erich & Sebastian Kämpf (2002). "The Impact of Intelligence and Institutional Improvements on Economic Growth". *Kyklos* 55(3) (August): 361–380

Appendix

A.1 Detailed Data Description

Educational attainment is the mean years of primary, secondary, and post-secondary education received by the population aged 15 and older, normalized for differential duration of education across countries. The (gross) *enrollment rate* is the share of pupils enrolled at the secondary level, regardless of age, relative to the theoretical age group for that level.

The *consumption-investment ratio* is total household and government consumption expenditure divided by gross fixed capital formation (gross of changes in the level of inventories), in constant 2000 U.S. dollars.

The *pupil-teacher ratio* is the number of pupils enrolled in primary school, divided by the number of primary school teachers. The additional school input is *public education expenditure*, which is the current and capital government spending on educational institutions (both public and private), education administration, as well as educational subsidies for private entities, such as households.

Kaufmann *et al.* (2007) collect governance data according to six dimensions: Voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption. As discussed in the text, the measure of *governance* that we employ for most specifications includes only the variable most likely to operate through human capital accumulation: government effectiveness. Estimates for this variable are assumed to be drawn from a normal distribution centered on zero with support $[-1, 1]$. We use the lagged effectiveness variable from the year 1996. For the fuller governance measure, we equally weight the 6 dimensions in the composite score to obtain an aggregate governance measure.

The two instruments we use for governance in the robustness section are common to those in Acemoglu *et al.* (2001) and Hall & Jones (1999): The mortality rates of early European settlers and the fraction of the population of European descent, specifically those speaking English and other European languages.

For additional controls introduced in the robustness section: *Trade* openness is taken to be net exports as a share of GDP, *geography* is the longitudinal distance from the equator, and *infrastructure* is proxied by road density, measured as kilometers of road per 100 square kilometer of land area.

We obtained fractionalization data from Alesina, Easterly, Devleeschauwer, Kurlat & Wacziarg (2003), democracy data from the Polity IV project Marshall & Jaggers (2005), and social capital data from the World Values Survey. *Ethnolinguistic fractionalization* is the sum of the ethnic and linguistic fractionalization measures, which in turn were computed as one minus the Herfindahl indices of the respective group shares in the population. The theoretical distribution has the range $[0, 2]$, with higher values indicating greater fractionalization. *Democracy* is a composite indicator of the competitiveness of executive recruitment and political participation, the openness of executive recruitment, and the strength of constraints on the chief executive; it has the integer range $[0, 10]$, with higher values indicating greater democracy. *Social capital* is a measure of trust in the society, which is calculated from the response to the question,

> "Generally speaking, would you say that most people can be trusted
> or that you need to be very careful in dealing with people?"

The indicator is binomial, distribution on support $[1, 2]$, with lower values indicating greater levels of trust. Following the literature, we assumed that trust was time-invariant, and so countries with more than one survey were collapsed into a single score by simple averaging.

For alternative variables used in the robustness section: *Parental authority* is the father's share of parental authority, which ranges from 0 (half share) to 1 (full). This was obtained from the OECD's Gender, Institutions, and Development database. *Ability* was calculated average national IQ estimates, adjusted to account for time differences as a result of the Flynn (2007) effect. This was due to Lynn & Vanhanen (2002). *Attainment* is the sum of the student performance in math and reading tests, adjusted for cross-country and cross-test comparability, from Altinok & Murseli (2007).

We used *adult schooling* as an alternative measure of family input; this is the mean schooling of the population aged 25 and over, and it serves as a proxy for parental education as a family input into the education process. Adult (youth) *literacy* is the percentage of the population aged 15 and older (aged 15–24) that is able to read and write a short, simple statement on their everyday life.

A.2 Bayesian Analysis of Exclusion Restriction

First, projections of the dependent variable Y/L, endogenous regressor H/L, and instrument G on the exogenous variables in the first stage, namely s_k and $(n+g+\delta)$. Second, residuals corresponding to these projections were then collected, and the variance of residuals corresponding to the instrument was normalized to one. Third, 10,000 draws were taken from the posterior distribution of π_2, for alternative values of ω. Finally, the 2.5th, 50th, and 97.5th percentiles of this distribution were computed, together with the interquantile ranges. The procedure is described in greater detail in Kraay (2008).

A.3 Additional Subsamples

We report additional panel regressions of subsamples of the data divided in two separate ways: (a) Geographic distribution, with countries groups into five broad regions: OECD, Latin America, Asia (to which we include South Asia), the Middle East, Eastern Europe, and Africa (Table A.1); (b) Income level, with countries grouped into high income (including both OECD and non-OECD countries), lower-middle and upper-middle income, and low income (Table A.2).

Table A.1: Panel regressions of GDP per capita, by region†

	OECD	L. America	Mideast	Asia	E. Europe	Africa
Investment	0.044	0.402	0.039	0.465	0.232	0.012
share	(0.13)	(1.96)	(0.07)	(0.20)**	(0.12)**	(0.04)
Net rate of	-0.027	0.638	-0.843	0.128	0.015	-0.243
depreciation	(0.15)	(1.89)	(0.45)*	(0.32)	(0.13)	(0.14)*
Human	-0.696	1.409	0.531	0.811	5.068	0.245
capital	(0.99)	(7.39)	(0.26)**	(0.26)***	(1.27)***	(0.18)
F	0.237	0.047	11.991***	9.617***	27.208***	4.835***
Anderson LR	6.121*	0.112	3.526	9.566**	6.042	9.508**
Cragg-Donald F	1.988	0.032	1.029	3.155	1.803	3.159
Hansen J	2.316	1.361	2.532	2.208	2.697	2.120
N	101	48	36	55	29	135

† Notes: Heteroskedasticity, cluster, and autocorrelation-robust (asymptotic) standard errors reported in parentheses. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.

Table A.2: Panel regressions of GDP per capita, by income level†

	Low	Lower middle	Upper middle	High
Investment	0.034	0.121	0.248	0.133
share	(0.06)	(0.13)	(0.10)**	(0.10)
Net rate of	0.128	-0.140	0.161	-0.173
depreciation	(0.22)	(0.25)	(0.10)*	(0.20)
Human	0.561	1.073	1.860	1.949
capital	(0.30)*	(0.30)***	(0.64)***	(1.15)*
F	3.689***	9.429***	4.786***	16.100***
Anderson LR	10.264***	12.344***	8.166***	4.915
Cragg-Donald F	3.423	4.168	2.668	1.598
Hansen J	1.327	1.967	6.138**	2.510
N	121	111	71	132

† Notes: Heteroskedasticity, cluster, and autocorrelation-robust (asymptotic) standard errors reported in parentheses. * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** indicates significance at 1 percent level.