Review of Development Economics, 20(2), 426–442, 2016
DOI:10.1111/rode.12236

Institutional Quality Mediates the Effect of Human Capital on Economic Performance

*Jonathon Adams-Kane and Jamus Jerome Lim**

Abstract

This paper considers the relationship between institutional quality, educational outcomes, and economic performance. More specifically, we seek to establish the linkages by which government effectiveness affects per capita income via its mediating impact on human capital formation. Our empirical approach adopts a two-stage strategy that estimates national-level educational production functions that include government effectiveness as a covariate, and uses these estimates as instruments for human capital in cross-country regressions of per capita income. Our results identify a significant and positive effect of human capital on per capita income levels, and partially resolves the inconsistency between macro- and micro-level studies of the effect of human capital on income. The results remain robust to alternative specifications, extension to a panel setting, subsamples of the data and fully endogenous institutions.

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 macroeconometric evidence. While microeconometric labor research using Mincer-style earnings functions generally find that human capital is one of the strongest predictors of lifetime income at the cross-country level (Peracchi, 2006), this intuitive result does not generally survive aggregation: while earlier studies did uncover a contribution of human capital to economic performance (Mankiw et al., 1992), later papers (Pritchett, 2001) have found either an insignificant or even *negative* relationship between human capital and income.

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 (Hanushek and Kimko, 2000), may suffer from systematic data deficiencies (Cohen and Soto, 2007), or arise from high rates of measurement error in cross-country education data (Soukiazis and Cravo, 2008). 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

*Lim: The World Bank, MSN MC2-204, 1818 H St NW, Washington DC, 20433, USA. Tel: +1-202-458-9120; Fax: +1-202-614-9120; E-mail: jlim@worldbank.org. Also affiliated to Santa Cruz Institute for International Economics. Adams-Kane: The Milken Institute, 1250 4th St, Santa Monica, CA, 90401, USA. 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, Gunilla Pettersson, Niklas Potrafke, the editor, Kwan Choi, and an anonymous referee for comments, as well as participants at the ISNIE, APEA, and SWPE meetings. 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; nor do they necessarily represent the views of the Milken Institute.

schools erode the productivity of the education sector (Rogers, 2008) and reduce the incentives for human capital accumulation (Gupta et al., 2001). This is an institutional failure that can spill over into growth outcomes. 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.

Complicating this basic story is the fact that empirical studies of the role of human capital in growth need to confront endogeneity in the former. While there is a strong theoretical basis for how human capital can drive growth in both neoclassical and endogenous models, there is also the possibility of reverse causality, possibly through a discount rate channel (Bils and Klenow, 2000). This endogeneity suggests that naïve attempts to measure the contribution of human capital will almost certainly encounter biases in their estimates.

Our empirical approach is a two-pronged one, designed to account for governance and measurement issues, while simultaneously addressing the endogeneity problem. Our first prong implements a two-stage estimation strategy that first regresses national-level educational production functions that include government effectiveness and inputs to schooling as covariates, then uses 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 seeks to address the endogeneity concerns that arise when using direct measures of education in 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. Our second prong comprises a suite of robustness checks that are designed to verify the quality of our instrument. These include tests for weak instrumentation, tests to rule out alternative channels whereby government effectiveness may potentially influence income, and fully endogenizing institutions and relying entirely on internal instruments via a system generalized method of moments (system GMM) estimator.

Our main results are supportive of the notion that human capital is central to cross-country incomes. Our benchmark specifications find that a 1% increase in human capital contributes 3.02–3.33% 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 that allow governance to be endogenous to income and/or human capital. These main results also follow when we alter our specification to exploit panel data, estimated via IV and system GMM.

We regard these findings to be of both academic and policy interest. Empirical studies of human capital have frequently struggled with 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 two papers closest in spirit to our own are the ones by Dias and Tebaldi (2012) and Hanushek and Kimko (2000). Like us, the first paper is interested in the relationship between institutions, human capital and economic performance; the theoretical analysis links institutions to human capital, but the econometric approach does not adopt a two-step procedure to estimating the mediating effect that governance plays on human capital, as we do here. The latter paper does use a similar two-step estimation procedure, but the first stage does not include our key conditioning variable of interest, institutional quality. Moreover, in contrast to both papers, our first stage is estimated in levels, rather than growth rates (which follows directly from our theoretical setup).

2. A Model of Human Capital and Governance

Our motivating theory is an augmented Solow growth model, with 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

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

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

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

where we further assume a Cobb–Douglas form and decreasing returns to inputs with $\gamma + \varepsilon < 1$. Innate ability is assumed to be distributed normally across countries at the global level such that there are no significant cross-country differences; thus, it is omitted at the aggregate level. Taking logarithms of (2) gives the (steady-state) amount of human capital per effective unit of labor:

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

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 $A_t = A_0 e^{gt}$ and $L_t = L_0 e^{nt}$, giving accumulation that follows

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

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

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. Steady-state levels of physical and human capital that follow are thus

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

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

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

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

3. Empirics of Income, Education and Institutions

Empirical Model

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

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

$$\ln\left[\frac{Y_{jt}}{L_{jt}}\right] = \pi_0 + \rho_j + \pi_1 \ln s_{k,jt} + \pi_2 \ln\left[\frac{H_{jt}}{L_{jt}}\right] - \pi_3 \ln(n+g+\delta) + \mathbf{X}_{jt}\mathbf{\Pi}_4 + v_{jt}, \tag{7}$$

where G_{it} is governance, $\mathbf{F}_{jt}$ and $\mathbf{S}_{jt}$ are vectors of family and school inputs to human capital production for country j at time t, respectively, H_{jt} is human capital, $s_{k,jt} = \frac{I_{jt}}{Y_{jt}}$ is the investment share of GDP, $(n + g + \delta) = n + 0.05$ is the net rate of depreciation of effective units of labor,[1] $\mathbf{X}_{it}$ is a vector of additional controls, Y_{it} is GDP, μ_i and ρ_i are time-invariant country fixed effects, and $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and $v_{it} \sim N(0, \sigma_v^2)$ 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 $\mathbf{X}_{it}$ 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 performance. In our benchmark specifications, we maintain parsimony with only one input for F and S; we relax this restriction in our robustness section. The benchmark estimates are cross-sectional in nature, so the t subscript above would not apply (although it would for our panel results).

Estimation and Identification Strategy

In our benchmark tests, we employ three main variables in our first-stage regressions. We contend that, of these three, two can be treated as plausibly exogenous and could thus function as instruments; the third may suffer from simultaneity concerns and so we consider a number of first-stage specifications that exclude this variable.

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 at the primary and secondary level is largely publicly provided.[2] If government effectiveness does matter to per capita incomes, there is a strong likelihood that it does so mainly through its mediating effect on the delivery of education.

There are two other main channels by which effective government may affect economic outcomes. The (ostensibly) most obvious channel is through policy, especially (but not limited to) macroeconomic policy. While this may be a plausible theoretical consideration, this seems to be less of an issue in practice. There is fairly abundant evidence that policy variables in general do not exert a systematic influence on economic growth, at least at the margin (Sala-i-Martin, 1997).[3]

The second channel is through public financial management. This should be understood to comprise two elements. The first is the quantity of public finances dedicated toward education, while the second is the quality of delivery of this public financing. Empirical work has found little evidence that public education expenditures (Sala-i-Martin et al., 2004) matters for growth. Insofar as the quality of financial management is involved, it would appear that while severe *mis*-management of public finances—in the form of corruption—have been found to affect growth directly, there is no clear reason why higher-quality public financial management should be expected to exert any first-order impact on growth, given that the quantity dimension appears to have so little impact.

In the analysis that follows, we perform a number of falsification tests that help rule out the possibility that these additional channels may invalidate the exclusion restriction. Finally, remaining simultaneity concerns are addressed by using a lagged specification of the effectiveness variable. Overall, we are reasonably confident that government effectiveness satisfies the exclusion restriction in the first stage. For completeness, however, we also provide several additional tests of the strength of this particular assumption when we discuss the benchmark results.

While an obvious candidate for household inputs is income per capita, it is essentially the same as the left-hand-side variable in the second stage regression and thus clearly not exogenous. Thus, the second instrument that we use is the consumption–investment (C/I) ratio, which acts as a proxy for family inputs into education.[4] To the extent that household educational expenditures are an investment good, the C/I ratio offers a plausibly exogenous instrument for family inputs that is not systematically related to per capita incomes.[5]

Our final variable is the pupil–teacher ratio, which is our proxy for school inputs (the two have an inverse relationship). We choose this variable, instead of other candidates, in part because of the strong case made for class size as a key determinant of schooling outcomes as a result of 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 reduced efficiency of the estimates owing to a poorer fit in the first stage. We navigate this tradeoff by reporting 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. In any case, we take steps to address this issue in our robustness section by including a range of control variables.

Estimation of the model is via two-stage least squares, using two-step 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 and 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.[6]

Data Description

Our cross-country macroeconomic data are drawn mainly from the World Bank's *World Development Indicators*. This includes the main controls: the investment share of output and the net rate of depreciation (which is the sum of the population growth rate and 0.05, as discussed in section 2). We supplement these with data from several other sources. Our primary measure of the human capital stock is the Barro and Lee (2001) dataset on educational attainment, which captures the average educational attainment of the population aged 15 and older. Our supplementary educational data—such as the pupil–teacher ratio and educational expenditures—are mainly from the UNESCO Institute for Statistics' *Global Education Statistics* database. Our primary governance data are from the *Worldwide Governance Indicators* (Kaufmann et al., 2011), which is both high quality and disaggregated into the subcomponents that we need. For reasons discussed in earlier, our preferred measure is government effectiveness, although we consider aggregate measures of governance as well.

The government effectiveness measure is central to our analysis and it is worthwhile describing it briefly here. The measure captures, *inter alia*, perceptions regarding the quality of public services and the quality of the civil service (Kaufmann et al., 2011). Key for our application is that this variable is a reasonably good proxy for the quality of educational service delivery, as distinct from other institutional quality measures—such as the rule of law or voice and accountability—that are likely to have a far weaker relationship with educational attainment, if any.

The benchmark sample comprises 64 developed and developing economies, for the year 2000. This year selection was dictated by data limitations: it is the only year when there is overlap between our preferred human capital measure and the governance measure used as an instrument. In a later subsection, we consider an alternative setup that comprises an unbalanced panel of 445 observations, for the years 1998, 2000, 2002–2006 (annual). Further details can be found in our working paper (Adams-Kane and Lim, 2014).

Main Findings

In Table 1 we report the main results of our benchmark model, which is a cross-section using 2000 data. 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.

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

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	−0.900	0.815	0.801	1.889	0.744	1.695
depreciation	(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: All variables were transformed to logarithmic form. 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. *,**,*** Indicate significance at 10%, 5% and 1% levels, respectively. For test statistics, the absence of asterisks is likewise indicative of statistical insignificance at these conventional levels.

school inputs and the consumption–investment ratio as a proxy for family inputs. Because of 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 physical capital is also consistent with the theoretical prior, but only marginally significant.[7] 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 identification ($F = 9.78$, $p = 0.00$).

The bottom half of column (B2) reports the corresponding first-stage results. While these estimates are of secondary interest, we note that in most of the specifications, these coefficients are statistically significant at conventional levels and yield either *a priori* expected signs. For example, since school resources are measured by the pupil–teacher ratio, an *increase* in this variable would be associated with *reduced* resources, which in turn tends to reduce student learning; hence the negative coefficient. Importantly, there is a positive relationship between improved governance and human capital. 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 is 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. (2011). 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)–(B6), respectively. The coefficient π_2 remains robust through these changes, although these are not directly comparable because of changes in the sample size resulting from differential data availability.

Taken together, the IV results reported in Table 1 suggest that a 1% increase in human capital contributes between 3.02% and 3.33% 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% to 1.10%. As is common for cross-country growth regressions, the large and significant constant term suggests that a substantial unexplained component remains. The broad qualitative message that emerges from this result—that institutions play an important mediating role in influencing how human capital affects macroeconomic performance—corroborates those of other papers that make broadly the same argument (Dias and Tebaldi, 2012).

These specifications also satisfy the primary diagnostic tests for instrument validity. We note that the Hansen *J* statistic 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 earlier). To formally test the validity of this important assumption, we exploit a recent procedure developed by Kraay (2012), which utilizes Bayesian inference to explicitly characterize the extent to which prior uncertainty about the assumption affects the posterior distribution of π_2.

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 values representing greater certainty that the exclusion restriction is valid. The support—for the 2.5th and 97.5th percentiles—is chosen to correspond to a 95% confidence interval; changes in the interquantile range are also reported.

Relative to the case where there is no prior uncertainty about the exclusion restriction ($\omega = \infty$), the support 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 1. 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.[8]

We perform one final set of tests for the exclusion restriction, based on the two other possible channels where government effectiveness might be expected to

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.

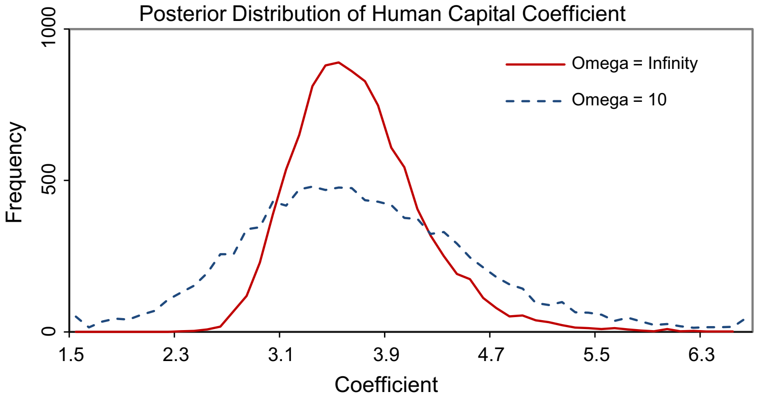

Figure 1. Posterior Distribution for Coefficient of Human Capital, with Alternative Assumptions about the Validity of the Exclusion Restriction
Notes*: Lower values of ω indicate greater prior uncertainty that the instrument satisfies the orthogonality condition. Even with high uncertainty, the posterior distribution of the coefficient does not include zero.*

operate. We perform a number of falsification tests by introducing a measure of either quality of macroeconomic policy or public financial management, alongside the human capital measure. For policy, we utilize the first principal component of an index of primary balances, gross debt and inflation, akin to the measure proposed in Burnside and Dollar (2000). For public financial management, we adopt an index of the quality of public investment management developed by Dabla-Norris et al. (2011).[9] By and large, these variables display low correlations with human capital (–0.20 for macro policy, 0.19 for structural policy and 0.19 for financial management), which alleviates concerns of multicollinearity. We find that, in regressions where we include both human capital and another alternative channel, the coefficient on the human capital variable remains significant, while those for either macroeconomic policy or financial management are insignificant (results available on request). We are thus fairly confident that these channels are not responsible for the effect of government effectiveness on income.

Robustness of the Benchmark

In the benchmark models, we did not introduce any additional controls to explain cross-country income per capita. Here, we allow **X** to include variables that the literature has identified as important. More specifically, we draw on a selection of the variables that Sala-i-Martin (1997) argues are robust empirical relations: the trade share of GDP, geographic location and infrastructure.[10] To this set we add some relatively more recent candidates in the empirical growth literature: ethnolinguistic fractionalization (Easterly and Levine, 1997), democratic development (Rivera-Batiz, 2002) and social capital (Knack and Keefer, 1997). These are reported in columns (R1)–(R6) of Table 3.[11] The significance of the human capital coefficient survives the inclusion of each of these additional controls.

We now proceed to consider alternative variables for and permutations of our exogenous instruments. In column (Z1) of Table 4, we consider an alternative way

Table 3. Regressions of GDP Per Capita with Additional Controls

	(R1)	*(R2)*	*(R3)*	*(R4)*	*(R5)*	*(R6)*
Investment share	0.931	0.752	0.663	0.879	1.527	1.195
	(0.51)*	(0.44)*	(0.74)	(0.47)*	(0.73)**	(0.47)**
Net rate of depreciation	0.853	0.997	0.779	0.935	1.303	0.444
	(1.00)	(0.88)	(1.13)	(1.10)	(1.56)	(0.89)
Human capital	3.160	2.901	2.992	3.203	4.076	2.646
	(0.49)***	(0.49)***	(0.54)***	(0.50)***	(0.77)***	(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	7.565	6.854	7.400	6.924	7.326
	(2.56)***	(2.11)***	(3.13)**	(2.74)***	(3.78)*	(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: All variables were transformed to logarithmic form. Huber–White (robust) standard errors reported in parentheses. *,**,***Indicate significance at 10%, 5% and 1% levels, respectively. For test statistics, the absence of asterisks is likewise indicative of statistical insignificance at these conventional levels.

to qualify family inputs in the education production function, using the notion that families with a greater share of parental authority invested in the mother—usually as a result of higher levels of education attained by them—are more likely to invest a greater share of family resources in education (Carneiro et al., 2013). This variable is used as an additional instrument to proxy for family inputs and while this weakens the instruments, human capital remains positive and significant.

Some authors have made a case for how the general intelligence quotient factor (Spearman's *g*) can affect growth (Jones and Schneider, 2006). Bearing in mind reservations over the use of *g* as a reliable indicator of general intelligence, we introduce a measure of intelligence, owing to Lynn and Vanhanen (2002), as a strong proxy for all resource inputs (so that the instrument set includes only IQ and government effectiveness). This is reported in column (Z2), and our results are largely unchanged.

In the specifications listed in Table 1, we shied away from using achievement data (in the form of test scores). By and large, the international comparability across different test types and time periods is suspect; 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 achievement data (Altinok and Murseli,

Table 4. Regressions of GDP Per Capita with Alternative Controls

	(Z1)	*(Z2)*	*(Z3)*	*(Z4)*	*(Z5)*
Investment share	1.636	0.550	0.013	−1.067	0.970
	(0.49)***	(0.23)**	(0.51)	(0.10)***	(0.42)**
Net rate of depreciation	1.672	0.821	4.464	−1.313	0.306
	(1.17)***	(0.65)	(1.07)***	(0.61)**	(0.82)
Human capital	3.556	3.242		1.788	3.073
	(0.48)***	(0.36)***		(0.20)***	(0.40)***
Alternative human capital			7.981		
			(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 share	0.761	0.972	0.895	−0.311	0.805
	(0.43)*	(0.37)***	(0.41)**	(0.47)	(0.49)*
Net rate of depreciation	0.853	0.997	0.779	−1.392	1.303
	(1.00)	(0.88)	(1.13)	(0.66)**	(1.56)
Human capital	3.070	3.139	2.538	0.707	1.897
	(0.46)***	(0.48)***	(0.75)***	(0.36)*	(1.14)*
Governance			0.203	1.654	0.320
			(0.22)	(0.30)***	(0.55)
Constant	6.886	7.493	7.096	3.179	5.700
	(2.20)***	(2.29)***	(1.98)***	(1.91)*	(1.79)***
Adj R^2	0.550	0.529	0.696	0.726	0.812
Anderson *LR*	31.628***	31.680***	9.507***	13.461***	1.091
Cragg–Donald *F*	9.108	9.127	3.096	4.508	0.245
Hansen *J*	1.907	1.933	1.521	0.034	4.395
N	64	64	64	23	64

Notes: All variables were transformed to logarithmic form. Huber–White (robust) standard errors reported in parentheses. *,**,*** Indicate significance at 10%, 5% and 1% levels respectively. For test statistics, the absence of asterisks is likewise indicative of statistical insignificance at these conventional levels.

2007) to examine how our results may change when we utilize a more accurate measure of human capital quality. The results are reported in column (Z3). Human capital remains significant and in this case its contribution more than doubles, so that a 1% increase in human capital leads to an almost 8% 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 tend to be important. 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). Although the results are once again similar, we note that the specification suffers from a small sample problem, which may limit inference.[12]

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 the most potentially problematic instrumental variable, government effectiveness, by including its contemporaneous value in the second stage. 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% level, while the coefficient for physical capital is also significant at the 5% 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 some limited confidence that the effects of good governance—at least when measured with government effectiveness—operate primarily through its mediating impact on human capital.[13]

In column (Z9) we use a measure of the pervasiveness of informal payments as an instrument for government effectiveness. There are several reasons why we choose not to use this instrument more extensively. First, the correlation between informal payments and both government effectiveness and human capital is very low ($\rho = -0.25$ and -0.24, respectively). Second, the sample size—even in the attenuated sample—is extremely small. Finally, the instrument is relatively weak. Nonetheless, we note that in this specification, human capital remains marginally significant, and government effectiveness is now positive and statistically significant (although we would caution against attributing too much to this finding owing to the weak instrument problem).

The fairly large literature on institutions and growth that emerged following the paper by Acemoğlu et al. (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, there may be questions over the quality of the settler mortality data (Albouy, 2012).

In any case, for comparability with the rest of the literature, we follow Acemoglu et al. (2001) and Hall and Jones (1999) and include in our instrument set instruments corresponding to the fraction of the population of European descent (we maintain as instruments family and school inputs). This is reported in column (Z10). As expected, the quality of the combined instrument set is suspect. The specification does not pass the underidentification test and the Cragg–Donald F statistic suggests that the instruments are extremely weak. Human capital does show up marginally significant and governance remains an insignificant predictor of income, but we heavily discount this result owing to poor test statistic performance.[14]

Additional Robustness Checks

Because of 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 5-year intervals from 1960 to 2000, the panel is unbalanced and consequently the 116-country sample has on average only four observations per country. 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 capital measure with data on enrollment rates. The panel with enrollment rates alone is much larger: 176 countries, with an average of 7 years. Results from this expanded dataset (documented in Adams-Kane and Lim, 2014) yields significant coefficient estimates for human capital that range from 1.503 to 2.183 and largely corroborate the findings of the cross-section estimates.

Another robustness check seeks to endogenize as many of the variables that we have used as possible; of particular interest 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. However, it is useful to raise an important caveat: the estimation method used here (system GMM) relies on the sequential exogeneity condition—the assumption that current innovations are uncorrelated with current and lagged values of the independent variables—for identification, there is the possibility that if this condition is violated, the internal instrument set may be compromised. Still, lacking obvious exogenous instruments for governance in particular, we proceed with this approach, keeping this caveat in mind. Results from this exercise (documented in Adams-Kane and Lim, 2014) yield coefficients on human capital that are significant across all the specifications, ranging from 0.816 to 1.651.

Finally, given the centrality of institutional differences, we perform one final set of checks to tease out the mechanism driving our results. We dissect the panel into subsamples corresponding to the following: (i) The subsamples above and below the median; (ii) Half a standard deviation above and below the mean; (iii) One standard deviation above and below the mean, all with respect to the broad institutional governance measure. In this case, the significant coefficients on human capital (documented in Adams-Kane and Lim, 2014) range from 0.674 to 2.497. Most notably, human capital appears to matter in institutional environments that are either relatively strong or relatively weak: countries that fall in the extremes of the institutional quality distribution face systematically different challenges in translating human capital investments into growth outcomes.

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 argue that since the effectiveness of government service delivery is central to improving learning outcomes and building human capital, the role of institutional governance in a country's growth process operates primarily though its intervening effect on human capital. Using a range of empirical strategies, we have taken this theory to the data and found support for this conjecture at both the cross-sectional and panel level.

It is also helpful to recognize that our findings are remarkably complementary to one other major explanation for the micro–macro paradox raised in the introduction: to the extent that it is educational quality, and not merely time spent in school, that matters for raising incomes (Hanushek and Kimko, 2000), our explanation likewise stresses the quality dimension from the perspective of governance mechanisms that safeguard educational delivery. From a policy perspective, our findings stress the critical importance of improving government effectiveness as a means of improving the mechanism of human capital accumulation.

References

Acemoglu, K. Daron, Simon Johnson and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91 (2001):1369–401.

Adams-Kane, Jonathon P. and Jamus Jerome Lim, "Institutional Quality Mediates the Effect of Human Capital on Economic Performance," World Bank Policy Research working paper 6792, Washington, DC (2014).

Albouy, David Y., "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment," *American Economic Review* 102 (2012):3059–76.

Altinok, Nadir and Hatidje Murseli, "International Database on Human Capital Quality," *Economics Letters* 96 (2007):237–44.

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

Barro, Robert J. and Jong-Wha Lee, "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers* 53 (2001):541–63.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review* 90 (2000):1160–83.

Burnside, Craig and David Dollar, "Aid, Policies, and Growth," *American Economic Review* 90 (2000):847–68.

Carneiro, Pedro, Costas Meghir and Matthias Parey, "Maternal Education, Home Environments, and the Development of Children and Adolescents," *Journal of the European Economic Association* 11 (2013):123–60.

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

Dabla-Norris, Era, Jim Brumby, Annette Kyobe, Zac Mills and Chris Papageorgiou, "Investing in Public Investment: An Index of Public Investment Efficiency," *Journal of Economic Growth* 17 (2011):235–66.

Dias, Joilson and Edinaldo Tebaldi, "Institutions, Human Capital, and Growth: The Institutional Mechanism," *Structural Change and Economic Dynamics* 23 (2012):300–12.

Djankov, Simeon, Rafael La Porta, Florencio López-de Silanes and Andrei Shleifer, "The Regulation of Entry," *Quarterly Journal of Economics* 117 (2001):1–37.

Easterly, William R. and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics* 112 (1997):1203–50.

Gupta, Sanjeev, Hamid R. Davoodi and Erwin R. Tiongson, "Corruption and the Provision of Health Care and Education Services," in Arvind K. Jain (ed.), *The Political Economy of Corruption*, London: Routledge (2001): 111–41.

Hall, Robert E. and Charles I. Jones, “Why do some Countries Produce so much more Output per Worker than Others?” *Quarterly Journal of Economics* 114 (1999):83–116.

Hanushek, Eric A. and Dennis D. Kimko, “Schooling, Labor-force Quality, and the Growth of Nations,” *American Economic Review* 90 (2000):1184–208.

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

Kaufmann, Daniel, Aart C. Kraay and Massimo Mastruzzi, “The Worldwide Governance Indicators: Methodology and Analytical Issues,” *Hague Journal on the Rule of Law* 3 (2011):220–46.

Knack, Stephen and Philip E. Keefer, “Does Social Capital Have an Economic Payoff? A Cross-country Investigation,” *Quarterly Journal of Economics* 112 (1997):1251–88.

Kraay, Aart C., “Instrumental Variables Regressions with Uncertain Exclusion Restrictions: A Bayesian Approach,” *Journal of Applied Econometrics* 27 (2012):108–28.

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

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

Mankiw, N. Gregory, David Romer and David N. Weil, “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics* 107 (1992):407–37.

Peracchi, Franco, “Educational Wage Premia and the Distribution of Earnings: An International Perspective,” in Eric A. Hanushek and Finis Welch (eds), *Handbook of the Economics of Education*, vol. 1, Amsterdam: Elsevier (2006): 189–254.

Pritchett, Lant H., “Where Has All the Education Gone?” *World Bank Economic Review* 15 (2001):367–91.

Rivera-Batiz, Francisco L., “Democracy, Governance, and Economic Growth: Theory and Evidence,” *Review of Development Economics* 6 (2002):225–47.

Rogers, Mark Llewellyn, “Directly Unproductive Schooling: How Country Characteristics Affect the Impact of Schooling on Growth,” *European Economic Review* 52 (2008):356–85.

Sala-i-Martin, Xavier X., “I Just Ran Two Million Regressions,” *American Economic Review* 87 (1997):178–83.

Sala-i-Martin, Xavier X., Gernot Doppelhofer and Ronald I. Miller, “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach,” *American Economic Review* 94 (2004):813–35.

Soukiazis, Elias and Túlio A. Cravo, “Human Capital and the Convergence Process among Countries,” *Review of Development Economics* 12 (2008):124–42.

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

Notes

1. 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.
2. Education expenditures typically account for about 14% of government expenditures, which is typically the largest single budget item (exempting social security in a small minority of countries).
3. This claim should be qualified. There is some evidence that very bad policy choices—such as financial repression or severe trade restrictions—may negatively affect country performance (Easterly and Levine, 1997). However, policies that can be more directly associated with government effectiveness—such as monetary and fiscal policy—tend to be insignificant in standard cross-country growth regressions.
4. Educational spending is generally classified as consumption expenditure by households; *ceteris paribus*, higher educational expenses would be associated with higher C/I ratios.

5. While consumption and investment are, independently, jointly determined with income, we are unaware of any theoretical models that posit a systematic link between the ratio of the two with income.
6. There are additional issues associated with the practical estimation of the augmented Solow model (Hall and Jones, 1999). We do not resolve these additional issues here, since our focus is on addressing the human capital puzzle, not testing the Solow model.
7. There is a concern that investment in physical capital may be endogenous to government effectiveness. While this may be important at the microeconomic level (Djankov et al., 2001), the channel does not appear to operate at the macroeconomic level, given the low correlation between the investment share and (lagged) government effectiveness ($\rho = 0.2$).
8. An important caveat to the tests is what the results would be if the distribution of priors were centered on a positive value. In this case, Kraay (2012) 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. We regard this theoretical possibility as unlikely, since the expected prior of the effect would have to be quite large—at least greater than the lowest 2.5th percentile value of 1.49—and estimates of the direct effect of governance on income yield coefficients that are comfortably below this value.
9. We also consider several alternative measures of each. For policy, for example, we used the World Bank's CPIA rating for macroeconomic management as well as structural policies, and for public finance, the World Bank's CPIA ratings of quality of budget management and efficiency of revenue mobilization.
10. We use 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 unaffected by these alternatives.
11. In addition to these controls, we considered a range of additional variables that could be of interest, such as regional dummies and government ideology, but our main were qualitatively unchanged.
12. Other permutations and combinations of these additional school inputs yielded similar significant coefficients for human capital, but typically did not satisfy the overidentification test.
13. Another possibility that would give rise to our result is that contemporaneous government effectiveness is strongly correlated with the other regressors included in the second stage. While this is not an issue for the investment share and net depreciation rate—the correlation coefficients are 0.29 and –0.34, respectively—this could be the case for human capital ($\rho = 0.70$). However, it is difficult to see how an increase in the current level of human capital accumulation can lead to a simultaneous increase in government effectiveness, given diffusion lags.
14. We also explored including the settler mortality instrument, with even more disastrous results: The instrument fails both the exclusion and relevance conditions and none of the variables in the second stage is statistically significant.