NBER WORKING PAPER SERIES

HEALTH AND ECONOMIC GROWTH:
RECONCILING THE MICRO AND MACRO EVIDENCE

David E. Bloom
David Canning
Rainer Kotschy
Klaus Prettner
Johannes J. Schünemann

Working Paper 26003
http://www.nber.org/papers/w26003

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2019

The authors would like to thank participants of the VfS Annual Congress 2018 and the European Meeting of the Econometric Society 2018 and Ana Abeliansky, Michael Burda, Alexander Khoury, Holger Strulik, and Uwe Sunde for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Health and Economic Growth: Reconciling the Micro and Macro Evidence
David E. Bloom, David Canning, Rainer Kotschy, Klaus Prettner, and Johannes J. Schünemann
NBER Working Paper No. 26003
June 2019
JEL No. I15,I25,J11,O11,O15

ABSTRACT

Economists use micro-based and macro-based approaches to assess the effects of health on economic growth. The micro-based approach tends to find smaller effects than the macro-based approach, thus presenting a micro-macro puzzle regarding the economic return on health. We reconcile these two strands of literature by showing that the point estimate of the macroeconomic effect of health is quantitatively close to that found by aggregating the microeconomic effects, when carefully specifying the estimation equations and controlling for spillovers of health at the aggregate level. Our results justify using the micro-based approach to estimate the direct economic benefits of health interventions.

David E. Bloom
Harvard T.H. Chan School of Public Health
Department of Global Health and Population
665 Huntington Ave.
Building 1, Suite 1202
Boston, MA 02115
and NBER
dbloom@hsph.harvard.edu

David Canning
Harvard School of Public Health
Department of Global Health and Population
665 Huntington Ave.
Boston, MA 02115
dcanning@hsph.harvard.edu

Rainer Kotschy
Ludwig-Maximillians-University Munich
Seminar for Population Economics Schackstr. 4 / I
Room 416 80539 Munich
rainer.kotschy@econ.lmu.de

Klaus Prettner
University of Hohenheim
Institute of Economics
Schloss Hohenheim 1d
70593 Stuttgart
Germany
Klaus.prettner@uni-hohenheim.de

Johannes J. Schünemann
University of Goettingen
Platz der Göttinger Sieben 3
37073 Göttingen
Germany
johannes.schuenemann@wiwi.uni-goettingen.de

1 Introduction

Health is an essential component of human capital that supports worker productivity by enhancing physical capacity and mental capabilities. Health improvements influence the pace of income growth through many pathways: Better health directly increases labor market participation and worker productivity (Strauss and Thomas, 1998; Bloom and Canning, 2000; Schultz, 2002; Bloom et al., 2003a); increasing life expectancy creates incentives to invest in education, innovation, and physical capital (Bloom et al., 2003b, 2007, 2014b; Cervellati and Sunde, 2013; Prettner, 2013); and better health, particularly that of women, reduces fertility and spurs an economic transition from a state of stagnating incomes toward sustained income growth (Galor and Weil, 2000; Galor, 2005, 2011; Cervellati and Sunde, 2005, 2011; Bloom et al., 2015). Hence, health and development correlate positively at the macroeconomic level, as the unconditional correlation between the survival probability from age 15 to 60 and (log) gross domestic product (GDP) per worker illustrates in Figure 1.

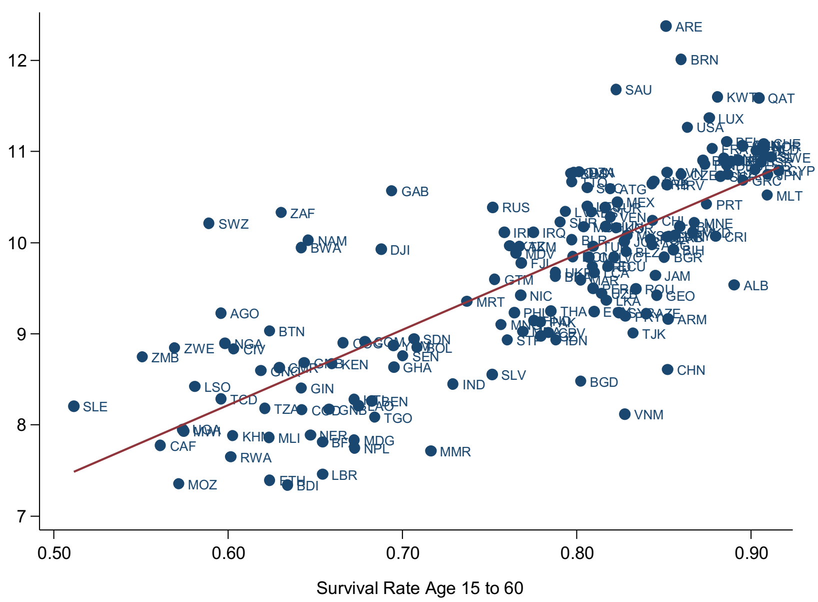

Figure 1: Unconditional Correlation: Health and Development

Sources: Feenstra et al. (2015) and United Nations (2017). Data are averaged over the time period 1970–2010.

In general, economists use two prominent methods to assess the effect of health on economic growth. The first aggregates the results of Mincerian wage regressions of the return on individual health to derive the macroeconomic effects of population health. The second relies on the estimation of a generalized aggregate production function that decomposes human capital into its components, including population health. While the overwhelming majority of studies based on both methods indicate a positive effect of health on economic growth, the size of the effect remains subject to intense debate. The former micro-based approach tends to find considerably smaller effects than the latter macro-based approach, thus presenting a micro-macro puzzle of the economic return on health.

This paper aims to reconcile both approaches by showing that the estimates based on microeconomic results are compatible with the effects derived from a well-specified macroeconomic analysis. To this end, we develop a production function model of economic growth, keeping our specification as close as possible to a generalized Mincerian wage equation as in Weil (2007). This permits us to compare our macro-level estimates and Weil's micro-level calibration directly. We address the potential endogeneity of the explanatory variables in several ways: First, we use predetermined measures of human capital inputs that are arguably exogenous in a dynamic cross-country panel when controlling for past economic development, time effects, and country-specific growth trends. Second, we exploit the demographic structure for an instrumental variables approach (see Kotschy and Sunde, 2018, who use this procedure to examine the implications of population aging and educational investment for macroeconomic performance). Third, we take advantage of the data's time structure to estimate dynamic panel general method of moments models. Although all three approaches rely on different identifying assumptions, they yield similar estimates of the macroeconomic return on health. Moreover, our results show that the micro-based and the macro-based estimates of the effects of health on economic development are quantitatively similar and, thus, consistent with each other. Hence, we provide a macro-based justification for using the micro-based approach to estimate the direct economic benefits of specific health interventions.

According to Weil (2007), a 10-percentage-point increase in adult survival rates translates into a 6.7-percent increase in labor productivity. Consequently, health differentials account for about 9.9 percent of the variation in output per worker across countries. Our macro-based analysis shows that a 10-percentage-point increase in adult survival rates is associated with a 9.1-percent increase in labor productivity. Weil's (2007) estimate falls well within the 95-percent confidence interval of our estimate, suggesting that the two models' results are compatible with each other. Because we include physical capital and education in our empirical framework, the resulting estimate is a measure of the direct productivity benefits of health as in Weil (2007). The estimated effect excludes the role of better health in increasing the incentives for investment, saving, and education, and its role in reducing fertility and spurring a takeoff toward sustained growth. Thus, the productivity benefits of health presented in this paper should be considered a conservative estimate of the overall economic benefits of health, including spillover effects.

Overall, our results and their consistency with the micro-based approach to estimating the return on health substantiate the claim that public health measures are an important lever for fostering economic development. These types of health investments could include vaccination programs, antibiotic distribution programs, and micronutrient supplementation schemes, which lead to large improvements in health outcomes for relatively low expenditures (World Bank, 1993; Commission on Macroeconomics and Health, 2001; Field et al., 2009; Luca et al., 2014).

The remainder of this paper is structured as follows. Section 2 reviews various approaches to measure the effect of health on economic performance. In Section 3, we derive the theoretical effect of health on output per worker from a human capital–augmented aggregate production function. In Section 4, we use these results to derive an empirical specification for estimating the influence of changes in population health on output growth. Section 5 describes the data, while Section 6 presents the empirical results. Finally, Section 7 concludes.

2 Literature Review

A common early empirical approach to examining the effect of health on economic growth involves regressing income per capita growth against initial level of health for a cross-sectional sample of countries, controlling for initial income and other factors believed to influence steady-state income (see, for example, Barro, 1991, 1997; Durlauf et al., 2005). Nearly all studies investigating economic growth that use this approach find a positive, significant, and sizable influence of initial population health—usually measured by life expectancy—on the subsequent pace of economic growth (see, for example, Barro and Sala-i-Martin, 2004; Easterly and Levine, 1997; Sachs et al., 1995; Sachs and Warner, 1997; Bhargava et al., 2001; Bloom et al., 2004). While the results for most other explanatory variables are not robust with respect to different specifications, Levine and Renelt (1992), Sala-i-Martin (1997), and Doppelhofer et al. (2004) find that initial population health is positively associated with subsequent growth in almost all permutations of explanatory variables they analyze. Hence, initial population health is one of the most robust predictors of subsequent economic growth.

More recent work analyzes the effects of health on economic growth via dynamic panel data regressions in the vein of Islam (1995), using the lagged dependent variable as one of the regressors to control for the convergence process.[2] These studies typically employ an exogenous instrument for health to isolate the causal channel running from better health to income growth (see, for example, Lorentzen et al., 2008; Aghion et al., 2011; Cervellati and Sunde, 2011; Bloom et al., 2014a). Acemoglu and Johnson (2007) is one of the few studies finding no evidence for a causal positive effect of health improvements on economic growth. The authors argue that increasing life expectancy raises population growth, which, in turn, increases capital dilution in the neoclassical growth model and therefore *reduces* income growth during the convergence process. They support this theory empirically using the global epidemiological revolution as an instrument for life expectancy. However, Aghion et al. (2011) and Bloom et al. (2014a) show that this result fails to hold when initial life expectancy is included in the regression. In addition, Cervellati and Sunde (2011) argue that Acemoglu and Johnson's (2007) results only hold for less-developed countries that have not yet undergone the demographic transition. In these countries, increasing life expectancy does indeed raise population growth and reduce income growth. For post–demographic transition countries, however, fertility declines with mortality. As a result, health improvements do not lead to an increase in population growth and capital dilution does not intensify. Splitting the sample into pre– and post–demographic transition countries, Cervellati and Sunde (2011) find that the effect of life expectancy on growth is positive for post–demographic transition countries and negative but insignificant for pre–demographic transition countries (see, for example, Hansen and Lønstrup, 2015, for a discussion).

Another way to assess the macroeconomic effect's size is by aggregating the microeconomic effects of health to infer the implications for aggregate income. For example, Fogel (1994, 1997) argues that much of British economic growth during 1780–1980 (about 0.33 percent per year) was due to increases in effective labor inputs that resulted from workers' better nutrition and improved health. More recently, the seminal works of Shastry and Weil (2003) and Weil (2007) employ an aggregate production function, in which the effects of health on productivity are calibrated from microeconomic wage regressions. These studies explain income by means of various health measures such as anemia, height, age at menarche, and the adult survival rate. The results of Shastry and Weil (2003) and Weil (2007)

[2] By construction, the fixed effects in such regressions correlate with the error term. Applying generalized method of moments estimators addresses this problem (see, for example, Arellano and Bond, 1991; Blundell and Bond, 1998; Judson and Owen, 1999).

suggest that health is an important form of human capital, but Weil (2007) argues that its effect on growth is smaller than that derived from macroeconomic cross-country regressions.

Somewhat in between these two types of studies are contributions that estimate the effects of health interventions at the macro level on income of individuals at the micro level. The most prominent example is the work by Bleakley (2007) analyzing the long-run benefits of campaigns to eradicate hookworm infections in the South of the United States. He finds that hookworm infections explain 22 percent of the income gap between the North and the South of the United States in 1900, which is consistent with macro-based and micro-based studies from a qualitative point of view. Other prominent studies show that the eradication or treatment of diseases, such as malaria, hookworm infections, and nutritional deficiencies, raises educational attainment, improves educational outcomes, and reduces fertility (Miguel and Kremer, 2004; Bleakley and Lange, 2009; Field et al., 2009). Bleakley (2010) discusses the problems that arise in estimating and interpreting the corresponding results.

3 Theoretical Framework

Assume that time $t = 1,2, \ldots$ evolves discretely, and consider an aggregate production function of the form

$$Y_t = A_t K_t^{\alpha} H_t^{1-\alpha}, \tag{1}$$

where Y_t denotes aggregate output, A_t represents total factor productivity (TFP), K_t is the physical capital stock, H_t describes the aggregate human capital stock, and α constitutes the elasticity of final output with respect to physical capital. The sum of individual levels of human capital $\nu_{j,t}$ of workers $j \in \{1,2, \ldots \mathcal{J}\}$ in the economy—that is, $H_t = \sum_j^{\mathcal{J}} \nu_{j,t}$—describes the aggregate stock of human capital. Expressing output in per worker units yields

$$y_t = A_t k_t^{\alpha} {\nu_t}^{1-\alpha} \tag{2}$$

with L_t being the size of the workforce and $y_t = Y_t/L_t$, $k_t = K_t/L_t$, and $\nu_t = H_t/L_t$. Alternatively, output can be expressed in per capita units as

$$\tilde{y}_t = \frac{Y_t}{N_t} = \frac{L_t}{N_t} A_t k_t^{\alpha} {\nu_t}^{1-\alpha}, \tag{3}$$

where N_t refers to the total population size.

In a competitive labor market, one unit of composite labor ν_t earns the wage w_t, which equals its marginal product:[3]

$$w_t = \frac{\partial y_t}{\partial \nu_t} = (1-\alpha)\frac{y_t}{\nu_t}.$$

Furthermore, individual human capital follows a generalized Mincerian wage equation along the lines of

[3] This holds under the assumption that a marginal change of individual human capital does not change the distribution of wages such that the marginal product of individual human capital and that of average human capital coincide.

Hall and Jones (1999), Bils and Klenow (2000), and Weil (2007). Hence, individual human capital $\nu_{j,t}$ follows an exponential function:

$$\nu_{j,t} = \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right), \tag{4}$$

where $h_{j,t}$ denotes the state of health, $s_{j,t}$ refers to educational attainment, $a_{j,t}$ describes work experience, ϕ_h is the semi-elasticity of human capital with respect to health, ϕ_s is the semi-elasticity of human capital with respect to educational attainment, and $\phi_{a,1}$ and $\phi_{a,2}$ refer to the semi-elasticities of experience and experience squared. We include the latter to capture the diminishing marginal contribution of experience to productivity.[4] Accordingly, a worker j with ν_j units of human capital earns a wage of

$$w_{j,t} = w_t \cdot \nu_{j,t} = w_t \cdot \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right). \tag{5}$$

This notation normalizes the effective labor input of a hypothetical worker without any health capital, education, and experience to unity. Meanwhile, workers with better health, higher education, or more experience are equivalent in productivity terms to a larger number of such baseline workers. Logarithmic wages at the individual level thus take the well-known Mincerian form:

$$\ln\left(w_{j,t}\right) = \ln(w_t) + \ln(\nu_{j,t}) = \ln(w_t) + \phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2. \tag{6}$$

Hence, the aggregate production function in (1) with our measure for human capital in (4) is consistent with wage equations used in the microeconomic literature.

The Mincerian wage form implies that the aggregate human capital stock in the economy is given by

$$H_t = \sum_j^{\mathcal{J}} \nu_{j,t} = \sum_j^{\mathcal{J}} \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right). \tag{7}$$

Accordingly, aggregating human capital requires raising individuals' health, educational attainment, and experience to the exponential power. This complication in the aggregation process vanishes if human capital and thus wages follow a log-normal distribution. In this case, the log of the average wage corresponds to the average of log wages plus one-half of the variance of log wages σ_t^2. Therefore, the log of human capital per worker simplifies to

$$\ln\left(\frac{H_t}{L_t}\right) = \ln\left(\frac{\sum_j^{\mathcal{J}} \nu_{j,t}}{L_t}\right) = \frac{\left[\sum_j^{\mathcal{J}} \ln\left(\nu_{j,t}\right)\right]}{L_t} + \frac{\sigma_t^2}{2}$$
$$= \frac{\sum_j^{\mathcal{J}} \phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2}{L_t} + \frac{\sigma_t^2}{2}$$

[4] Conceptually, $h_{j,t}$, $s_{j,t}$, and $a_{j,t}$ need not represent all aspects of health, educational attainment, and experience, only those that are relevant for the production of final output.

$$= \phi_h h_t + \phi_s s_t + \phi_{a,1} a_t + \phi_{a,2} a_t^2 + \frac{\sigma_t^2}{2}. \tag{8}$$

Intuitively, a marginally better health status (for example, an increase in the adult survival rate by 1 percentage point) raises a worker's productivity and wages by $100 \cdot \phi_h$ percent. Analogously, additional marginal investment in education (for example, one year of schooling) raises a worker's productivity and wages by $100 \cdot \phi_s$ percent. This effect's absolute size is larger for highly educated high-wage earners than it is for poorly educated low-wage workers. Moreover, an extra year of education for a highly educated worker also represents a greater investment, because the worker must forgo a higher wage for the extra schooling.

4 Empirical Framework

Suppose the production function in (2) applies to $i = 1, \dots, \mathrm{I}$ countries. Taking the logarithm of the production function and using the result from Equation (8), the log of production per worker is given by

$$\ln(y_{i,t}) = \ln(A_{i,t}) + \alpha \ln(k_{i,t}) + (1-\alpha)\left(\phi_h h_{i,t} + \phi_s s_{i,t} + \phi_{a,1} a_{i,t} + \phi_{a,2} a_{i,t}^2 + \frac{\sigma_{i,t}^2}{2}\right). \tag{9}$$

Using rates of return to calibrate the coefficient on education, ϕ_s, suggests a parameter value of 0.09–0.15 (Psacharopoulos, 1994; Bils and Klenow, 2000; Psacharopoulos and Patrinos, 2004, 2018). Regarding the elasticity of output with respect to capital, α, economists generally agree on values of around one-third (see, for example, Hall and Jones, 1999).

Similarly, Heckman and Klenow (1997) and Krueger and Lindahl (2001) derive a formula that estimates the macroeconomic effects of schooling using an aggregated version of a Mincerian wage equation. The major difference between these formulations is that education level's effect on output in their formulation is expressed as ϕ_s, whereas in our approach the effect of schooling is $(1-\alpha)\phi_s$. This difference arises because they assume the cross-country differences and changes in the intercepts in (6) are random and assign them to the error term in the regression. With our production function, increases in schooling increase the aggregate level of human capital and labor equivalent inputs in the economy and depress the wage paid per equivalent worker.

Equation (9) describes aggregate production as an identity that could be estimated directly, if all right-hand-side variables were available. In practice, however, the level of total factor productivity in country i at time t, $\ln(A_{i,t})$, is not observed. Several approaches can address this problem. We follow Bloom et al. (2004) and model total factor productivity as a diffusion process across countries, which allows for the possibility of long-run differences in TFP even after the diffusion is complete. Specifically, let the change in TFP be given by

$$\Delta \ln(A_{it}) = \lambda\left[\ln(A_{i,t}^*) - \ln(A_{i,t-1})\right] + \varepsilon_{i,t}, \tag{10}$$

where $\varepsilon_{i,t}$ constitutes an idiosyncratic shock. Each country has a period-specific upper bound, given by $\ln(A_{i,t}^*)$. A country's total factor productivity adjusts toward this bound at rate λ. We assume this upper bound depends on country characteristics $x_{i,t}$ and on the worldwide technology frontier μ_t. Moreover, schooling in previous periods may facilitate the diffusion and adoption of existing technologies (Nelson and Phelps, 1966) or spur novel innovation (Romer, 1990; Strulik et al., 2013). Hence, lagged schooling, $s_{i,t-1}$, constitutes another determinant of potential TFP (see also Cuaresma et al., 2014). Neglecting one of these channels might bias the empirical estimates, as Sunde and Vischer's (2015) results indicate. Because technological gaps are not directly observed, we follow Baumol (1986) and use lagged output per worker as a proxy (see also Fagerberg, 1994; Dowrick and Rogers, 2002). Hence, growth of total factor productivity reads

$$\Delta\ln(A_{it}) = \lambda\left[\mu_t + x_{i,t}{}'\Theta + \rho s_{i,t-1} - \ln(y_{i,t-1})\right] + \varepsilon_{i,t}. \tag{11}$$

Alternatively, a richer model derives the log of lagged total factor productivity $\ln(A_{i,t-1})$ directly from the production function such that

$$\begin{aligned}\Delta\ln\left(A_{i,t}\right) = {}& \lambda\left[\mu_t + x'_{i,t}\Theta + \rho s_{i,t-1} - \ln\left(y_{i,t-1}\right) + \alpha\ln\left(k_{i,t-1}\right)\right] \\ & + \lambda(1-\alpha)\left(\phi_h h_{i,t-1} + \phi_s s_{i,t-1} + \phi_{a,1} a_{i,t-1} + \phi_{a,2} a^2_{i,t-1} + \frac{\sigma^2_{i,t-1}}{2}\right) + \varepsilon_{i,t}.\end{aligned} \tag{12}$$

This slightly more comprehensive modeling approach, however, suffers from the disadvantage that including additional highly correlated explanatory variables inflates the estimated standard errors without providing additional insights into the parameters of interest. We provide estimates herein for both models and show that they are qualitatively and quantitatively similar.

Related research suggests several variables $x_{i,t}$ that may affect the TFP level in the long run. For example, Hall and Jones (1999) argue that institutions and "social infrastructure" affect productivity, while Gallup et al. (1999) emphasize the role of geography. Our empirical work experiments with several potential variables to control for these influences.

First-differencing (9) and inserting (11) provides the empirical estimation equation:

$$\begin{aligned}\Delta\ln\left(y_{i,t}\right) = {}& \lambda\left[\mu_t + x'_{i,t}\Theta + \rho s_{i,t-1} - \ln\left(y_{i,t-1}\right)\right] + \alpha\Delta\ln\left(k_{i,t}\right) \\ & + (1-\alpha)\left(\phi_h\Delta(h_{i,t}) + \phi_s\Delta(s_{i,t}) + \phi_{a,1}\Delta(a_{i,t}) + \phi_{a,2}\Delta(a^2_{i,t}) + \frac{\Delta(\sigma^2_{i,t})}{2}\right) + \varepsilon_{i,t}.\end{aligned} \tag{13}$$

De la Fuente and Domenech (2001) and Bloom et al. (2004) use this approach to model TFP diffusion in cross-country production function studies. It is formally equivalent to the autoregressive TFP model that Griliches and Mairesse (1998) and Blundell and Bond (2000) use in their studies of the production function based on firm-level data.

According to the specification in (13), output growth can be decomposed into three components. The first is a catch-up term capturing the reduction of the technological gap to the leading countries in each time period, such that the country converges to its TFP upper bound at the rate λ. The second is growth of the input factors capital, health, schooling, and experience. The third component is an idiosyncratic shock to the country's total factor productivity $\varepsilon_{i,t}$.[5]

Equation (13) represents a model of conditional convergence, in which the speed of convergence λ describes the rate at which gaps in total factor productivity close. Therefore, this approach contrasts with models that take TFP differentials across countries to be fixed, such as those of Mankiw et al. (1992) and Islam (1995). The speed of convergence in these models depends on the time that capital stocks take to reach their steady-state levels given fixed investment rates. By including the growth rates of factor inputs directly in Equation (13), we can identify the catch-up term—that is, the effect of the gap between actual and steady-state output, given current input levels—as the impact of a TFP gap.

In the special case of no technological diffusion ($\lambda = 0$), the lagged level terms in (13) disappear. Hence, our approach encompasses the estimation of a production function in first differences, as Krueger and Lindahl (2001) and Pritchett (2001) advocate. Moreover, we can test if this restriction holds. Taking first differences nets out any fixed effects on TFP. Therefore, testing whether $\lambda = 0$ examines the plausibility of TFP differentials remaining constant or, alternatively, narrowing over time because of technological diffusion. Our model also encompasses the special case in which technological diffusion occurs, but the steady-state level of TFP is the same in every country. We can test this by examining whether the country-specific variables $x_{i,t}$ have zero coefficients.

When estimating Equation (13), we face the possibility that contemporaneous growth rates of factor inputs are endogenous and responsive to the current TFP shock $\varepsilon_{i,t}$. We address this concern in several ways: First, we use predetermined measures of health and education inputs that are arguably exogenous given the approximation of TFP growth rates, which controls for productivity gains that are due to past changes in input factors, past technology shocks, and convergence to the technological frontier. This estimation strategy thus corresponds to an ordinary least squares (OLS) model. Second, we exploit the demographic structure to obtain plausibly exogenous instrumental variables (IV) for health and education inputs (see Kotschy and Sunde, 2018). Specifically, inflows from young-age cohorts at the lower end and outflows from old-age cohorts at the upper end of the working-age population determine changes in overall health and educational attainment of the working-age population. Hence, one can use the lagged levels of health and educational attainment for the age cohorts that will enter or leave the working-age population in the next period as an instrumental variable for the contemporaneous growth rate of the corresponding factor input. These instruments rely on variation in health and education inputs predating the estimation period, and they are plausibly exogenous given the approximation of TFP growth rates, which controls for past economic performance. Third, we take advantage of the data's time structure to estimate dynamic panel general method of moments (GMM) models, thereby eliminating potential spurious correlations that arise mechanically through a link between the lagged dependent variable and the error term. In particular, this approach instruments potentially endogenous regressors by their lagged values. We also estimate a full instrumentation specification that combines the instrumental variables with the GMM approach.

Each of these identification strategies is compatible with lagged TFP levels and expected TFP growth—the catch-up term in Equation (13)—affecting previous input decisions (for example, Bils and Klenow, 2000, suggest that schooling decisions depend on expected economic growth). The argument that

[5] We could allow the shock to grow over time to have a common component across countries, such as a worldwide oil or interest rate shock. Such a shock, however, would be collinear to changes in the worldwide productivity frontier captured by the time effects and would thus not affect any of our results.

lagged input levels are uncorrelated with future shocks to TFP is the rationale for estimating Equation (13) instead of the level relationship in Equation (9).

Finally, we also include fixed effects in comprehensive OLS and IV specifications to account for the possibility of country-specific growth trends driven by unobserved heterogeneity with a persistent effect on TFP.[6] These specifications are conservative, as they identify the parameters of interest solely based on deviations of countries' growth rates from their country-specific growth trends, thereby considerably confining the set of potential confounding factors. However, this conservatism implies a loss of precision in the estimates and may, moreover, increase the risk of underestimating the impact of factor inputs if there is measurement error in the independent variables (see Hauk and Wacziarg, 2009). Both of these factors complicate drawing inferences based on the fixed-effects results. We nonetheless report these results to probe the robustness of our baseline results.

5 Data

We construct an unbalanced panel of 116 countries observed every five years from 1970 to 2010. Data on real output and physical capital, both in per worker units, are obtained from the Penn World Tables by Feenstra et al. (2015).

Measures for health inputs are obtained from United Nations (2017). We use adult survival rates, which measure the probability of surviving from age 15 to 60. Conceptually, this measure may relate more closely to adult health and worker productivity than life expectancy, which is sensitive to changes in infant mortality rates. Adult survival rates, however, are only a proxy for the health of the workforce, because they measure mortality rates rather than morbidity. Our main reason for using adult survival rates is that it allows us to compare our results directly with those of Shastry and Weil (2003) and Weil (2007). We additionally report results for life expectancy in the robustness section.

Following the Mincerian approach, educational input is proxied by years of schooling in the working-age population. To this end, we exploit measures of secondary and total schooling from Barro and Lee (2013) for the population above age 15. We combine age-specific years of schooling with population shares to construct average years of schooling for the working-age population, which we define as age 15 to 60 to match our measure of aggregate health.

We construct aggregate experience as the median age of the population obtained from United Nations (2017), net of an intercept of six years corresponding to early childhood. Moreover, we deduct compulsory schooling years, taken from UNESCO (1997) and UNESCO (2017), to account for differences in the age of workforce entry across countries. This correction is necessary, because countries with higher life expectancy and older populations tend to have later workforce entry due to longer schooling. As experience enters the regression framework in differences, this measure takes up variation from changes in median age and compulsory schooling following educational reforms.[7]

To control for the effect of wage inequality, we use the disposable income Gini coefficient after

---

[6] Note that dynamic panel GMM estimators automatically remove fixed effects by design.

[7] For certain countries, the statistical yearbooks report values for specific regions. Moreover, some countries' educational systems allow for different categorizations, such that alternative figures are conceivable. We correct for these fluctuations and code less varying values in the case of doubt. This procedure tends to render the measure for experience less informative and thus increases the corresponding standard errors. Table A8 in the Appendix contains a complete list of coding decisions. Because we use only changes in experience over time, measurement error in compulsory schooling levels poses no threat to our identification.

taxation and transfers by Solt (2016b). These data provide standardized Gini coefficients that are comparable across countries and over time.

Finally, we include country-specific variables that may affect long-run TFP levels. These include an indicator for the quality of economic institutions from Gwartney et al. (2017), a measure for the value added by the agricultural sector from World Bank (2017) to control for structural change, the percentage of land area in the tropics by Gallup et al. (1999) to control for geographical factors that may affect productivity and trading opportunities, and a set of regional dummies.[8] Among these control variables, institutional quality is particularly important, as it may simultaneously raise income growth, health outcomes, and education inputs (Weil, 2014). Because the data on institutional quality are available only from 1970 onward, they restrict the estimation sample to the time period 1970–2010. Table A1 in the Appendix reports descriptive statistics for the estimation sample.

6 Results

6.1 Ordinary Least Squares Results

Table 1 presents the estimation results based on OLS and fixed-effects regressions. We proxy education by average years of secondary schooling, which accounts for the largest changes in educational attainment over the estimation period and which provides the most precise estimates for the return on education. Column (1) reports coefficient estimates of a parsimonious specification of our empirical model in Equation (13), including lagged educational attainment but omitting any additional controls. The point estimates show the sign expected from theory. Lagged per worker GDP is negative, implying conditional convergence as predicted by the neoclassical growth literature (Solow, 1956; Cass, 1965; Diamond, 1965) and as established empirically.[9] Capital accumulation positively relates to economic growth, which again conforms to the growth literature's results. Changes in the aggregate human capital stock positively affect productivity per worker: The coefficients for changes in population health and average years of schooling both show a positive sign and differ significantly from zero at the 1-percent level. Hence, health and education both constitute important dimensions of human capital. The opposing signs for $\hat{\phi}_{a,1}$ and $\hat{\phi}_{a,2}$ suggest a hump-shaped effect of average experience on growth of output per worker, which is consistent with the standard Mincerian framework; though only the coefficient of squared experience is (marginally) significant in this specification. Because experience varies strongly across individuals but very little across countries, however, obtaining a precise estimate of the effect of worker experience in macroeconomic models is difficult (Bloom et al., 2004).

[8] We also experimented with further indicators for landlocked countries (Gallup et al., 1999) and controls for ethnic fractionalization and polarization by Alesina et al. (2003) and Reynal-Querol and Montalvo (2005). Given the set of other controls, however, these variables did not explain much of the remaining variation.

[9] See, for example, Barro (1991, 1997), Sala-i-Martin (1997), and Doppelhofer et al. (2004) in cross-section regressions and Islam (1995), Caselli et al. (1996), and Brückner (2013) in panel data settings. For interesting surveys and critical remarks on the literature, see Durlauf et al. (2005) and Eberhardt and Teal (2011).

Table 1: Return of Health and Education to Productivity: OLS Results

	Baseline Configuration				Including Country-Specific Growth Trends			
	No Controls (1)	Adding Controls (2)	Adding Gini (3)	Lagged Controls (4)	No Controls (5)	Adding Controls (6)	Adding Gini (7)	Lagged Controls (8)
$\ln(y_{i,t-1})$	-0.045***	-0.15***	-0.18***	-0.20***	-0.31***	-0.34***	-0.39***	-0.44***
	(0.0097)	(0.017)	(0.023)	(0.022)	(0.053)	(0.052)	(0.046)	(0.069)
$\Delta \ln(k_{i,t})$	0.52***	0.35***	0.23***	0.48***	0.48***	0.29***	0.085	0.44***
	(0.054)	(0.063)	(0.067)	(0.075)	(0.095)	(0.10)	(0.082)	(0.11)
$\Delta(h_{i,t})$	1.12***	0.59**	0.74**	0.68**	1.01***	0.64**	0.55	0.81**
	(0.30)	(0.29)	(0.35)	(0.28)	(0.28)	(0.27)	(0.38)	(0.32)
$\Delta(s_{i,t})$	0.099***	0.075**	0.063**	0.063**	0.11**	0.091**	0.033	0.080*
	(0.032)	(0.031)	(0.027)	(0.029)	(0.049)	(0.046)	(0.032)	(0.040)
$\Delta(a_{i,t})$	0.011	0.0075	0.0089	0.0090	-0.0095	-0.0100	0.0070	-0.0051
	(0.0085)	(0.0088)	(0.010)	(0.0085)	(0.0100)	(0.0095)	(0.0081)	(0.0094)
$\Delta(a^2_{i,t})$	-0.00069*	-0.00057	-0.00058	-0.00064*	0.00026	0.00024	-0.00023	0.000083
	(0.00036)	(0.00036)	(0.00039)	(0.00033)	(0.00038)	(0.00036)	(0.00035)	(0.00036)
$\Delta(\sigma^2_{i,t})$	—	—	-0.60	—	—	—	-0.70*	—
	—	—	(0.45)	—	—	—	(0.42)	—
R^2	0.29	0.38	—	0.41	0.31	0.38	—	0.41
Countries	116	116	109	116	116	116	109	116
Observations	613	613	461	613	613	613	461	613
Controls	—	✓	✓	✓	—	✓	✓	✓
Fixed Effects	—	—	—	—	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$ and a full set of time effects. Columns (2), (3), (4), and (6), (7), (8) add further controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, and a full set of regional dummies. Columns (5) to (8) allow for country-specific growth trends by including fixed effects. The comprehensive specifications in Columns (4) and (8) additionally control for the first lag of physical capital per worker, population health, experience, and experience squared. The standard errors in all specifications are clustered at the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

The specification used in Column (2) is our baseline specification and includes controls for the quality of economic institutions, the value added by the agricultural sector (as a proxy for structural change), the percentage of land area in the tropics, and the set of regional dummies. Adding these controls increases the model's explanatory power, as reflected by an increase in R^2, and slightly improves the precision of the point estimates. Quantitatively, the computed parameters reduce in magnitude compared with the parsimonious specification, but the qualitative results remain unchanged. In particular, the estimates still indicate a positive and significant effect of changes in average health and education on output per worker. Interestingly, the reduction in magnitude of the return on health is almost entirely due to the inclusion of institutional quality. This confirms recent evidence by Weil (2014, 2017), who finds that institutional differences account for a considerable portion of the cross-country correlation between income and health. Nevertheless, our results also show that even after controlling for institutional differences, significant scope exists for a positive effect of health on output per worker.

In Column (3), we augment the specification in Column (2) by adding the Gini coefficient to approximate the variance of log wages. For reasons of data availability, the estimation sample shrinks to 461 observations.[10] The qualitative results again remain unchanged, while the estimated parameters do not change considerably in magnitude. The computed parameter for the disposable income Gini coefficient is negative and insignificant. Finally, Column (4) presents the results for a comprehensive model with additional lagged controls, which derives lagged TFP directly from the production function according to Equation (12). The results conform quantitatively and qualitatively to the results in Columns (2) and (3).

In Columns (5) to (8), we include country-fixed effects to allow for country-specific growth trends that take up unobserved heterogeneity with a persistent effect on TFP. These specifications substantially restrict the potential of omitted variables to bias the outcomes of interest and thus serve as a specification test for the models without fixed effects. Qualitatively, the results remain unchanged. Quantitatively, the estimated return on health does not change considerably and remains significant in all but one specification, in which we use the reduced sample to control for wage ineqality. Also the estimated return on education takes similar values as in Columns (1) to (4). Overall, including fixed effects slightly increases the standard errors of most coefficients without considerably improving the explanatory power of the models including additional controls. Our result of a positive and significant return on health is, therefore, not driven by country-fixed unobservables.

Table 2 compares the results of the baseline specification in Column (2) with those of the literature. Weil (2007) derives the macroeconomic return on health using a set of microeconomic estimates of the return on height from childhood inputs, twin studies, and long-run historical data. According to his baseline calibration, an increase in adult survival rates of 0.1—or 10 percentage points—raises labor productivity by 6.7 percent. To obtain this figure, he uses the two lowest but arguably best identified estimates, stemming from twin studies in the developed world. At the same time, these estimates are conservative, "[i]f nutrition primarily affects physical capabilities and if these capabilities are less important in rich than in poor countries […]" (Weil 2007, p. 1288). Averaging over all estimates and study types instead would suggest a larger effect on labor productivity of 13 percent. We follow Weil's (2007) approach and set our target to 6.7 percent; though, we deem values slightly above this target still to be plausible. Our estimates are quantitatively similar to this target and indicate that an increase in adult survival rates of 0.1 translates into a 9.1-percent increase in labor productivity.[11] Apart from that, the target serves as a conservative benchmark, the slightly larger point estimate from the macro-based approach might also be due to spillover effects that the micro-based approach omits by design. The 95-percent confidence interval of our estimate ranges from 0.47 to 17.7 percent and thus includes Weil's (2007) estimate. Consequently, our macro results are consistent with the micro results, reconciling the micro-based and macro-based approaches to estimating the effect of health on income growth.

---

[10] To increase data availability, Solt (2016b) uses imputation procedures to reduce the number of missing values in the data set. Because this procedure may understate the uncertainty in the data and thus lead to downward-biased standard errors, we conduct a standard-error adjustment as Solt (2016a) suggests. In particular, we estimate the specification in Column (4) for 100 potential realizations of the Gini coefficient and compute the final estimates as the average overall individual results. For details, see Solt (2016a).

[11] To obtain the figures for ϕ_h and ϕ_s in Equation (13), divide the estimates in Table 1 by $(1-\alpha)$.

Table 2: Comparison Between Our Estimates and the Literature

Variable	Point Estimate	Confidence Interval	Target
γ	< 0		< 0
α	0.35	(0.23–0.47)	0.3–0.4
ϕ_h	9.1%	(0.47%–17.72%)	6.7%
ϕ_s	11.4%	(2.6%–20.26%)	10.5%
$\phi_{a,1}$	> 0		> 0
$\phi_{a,2}$	< 0		< 0

Our coefficient estimate for changes in physical capital α is 0.35. This is in line with empirical estimates of the elasticity of output with respect to physical capital, which fall around 0.3 to 0.4 (Hall and Jones, 1999). By dividing our estimate of education by $(1-\alpha)$, we estimate a return on secondary schooling ϕ_s of 11.4 percent. This value is consistent with the estimates of the studies reviewed in Psacharopoulos and Patrinos (2018), where the return on secondary schooling over all studies and all countries in their supplementary file averages to 10.5 percent. Moreover, our point estimate of 11.4 percent falls within the range of the most plausible estimates in the literature (see Psacharopoulos, 1994; Bils and Klenow, 2000; Psacharopoulos and Patrinos, 2004, 2018). Finally, the signs of our coefficient estimates on the lagged dependent variable γ and our experience measures $\phi_{a,1}$ and $\phi_{a,2}$ accord with previous findings on conditional convergence and positive but diminishing returns to experience.

6.2 Instrumental Variables Results

Another way to address the potential endogeneity of health and education in the empirical model is to exploit the demographic structure of the working-age population for an instrumental variables approach. This identification strategy adapts and extends the approach in work by Kotschy and Sunde (2018). It relies on four observations. First, the demographic structure of the working-age population follows very predictable patterns. For example, individuals who were born in 1965 are 20 years old in 1985 and 25 years old in 1990. Second, the changes in health and education of the working-age population are to great extent determined by the in- and outflows of young- and old-age cohorts at the lower and upper ends of the working-age population. Note that migration is also less of a concern for these age groups than it is for prime-age workers. Third, childhood health is predictive of adult health (see, for example, Case et al., 2005; Currie et al., 2010). Fourth, childhood health and education decisions predate the estimation period, such that they are plausibly exogenous in a model that also captures past economic performance. To predict contemporaneous changes in aggregate health and education of the working-age population we can, therefore, use lagged childhood health and lagged education of the age cohorts that move into and out of the working-age population. Figure A1 in the Appendix illustrates this strategy.

Specifically, we instrument contemporaneous changes in working-age health by lagged childhood health of the cohort aged 10–14 years, measured in terms of child mortality at time of birth, as this cohort enters the working-age population in the subsequent period. Optimally, we would also like to capture the outflow of aggregate health by using lagged health of the cohort aged 55–59 years; however, a lack of data on child mortality rates from before the world wars prohibits us from doing so. Analogously, we instrument contemporaneous changes in education by average years of secondary schooling of the cohort aged 55–59 years in the previous period, as this cohort leaves the working-age population in the subsequent period. At the same time, we refrain from using the inflow of schooling for young-age cohorts, because individuals might anticipate future economic growth and thus increase educational attainment.

Table 3 displays the first-stage results. All specifications include past economic development, changes in input factors, time effects, and the set of baseline controls. Even-numbered columns account for country-specific growth trends by adding fixed effects. Columns (1) and (2) instrument changes in the adult survival rate by the lagged child mortality rate of the cohort aged 10–14 years, treating education inputs as exogenous. Columns (3) and (4) instrument changes in working-age education by lagged schooling of the cohort aged 55–59, treating health inputs as exogenous. The specifications in Columns (5) and (6) instrument changes in both health and education inputs. Because two endogenous variables are present, these specifications require two first-stage regressions: one for each variable. Throughout all specifications, lagged child mortality significantly predicts changes in adult survival rates, while lagged schooling of old-age workers significantly predicts changes in working-age education. In particular, a 1-percentage point higher child mortality rate of the cohort entering the working-age population is statistically associated with a roughly 0.2 percentage points higher subsequent increase in adult survival. This correlation is consistent with convergence in health technologies across countries. At the same time, one additional year of secondary schooling of the cohort leaving the working-age population is statistically associated with a roughly 0.1 lower subsequent increase in average secondary education of the working-age population. The computed coefficients do not differ considerably across specifications, indicating that invariant, country-specific heterogeneity does not pose a problem. In fact, including fixed effects does not improve the models' fit at all, as reductions in the models' R^2s show. The first-stage *F*-statistic takes values above 10 in all specifications without fixed effects, thereby ruling out weak instruments. Note that for models with two instruments, the combined *F*-statistic is relevant, because there are two equations. For the fixed-effects model in Column (6), the first-stage correlations between the instruments and the endogenous variables are not sufficiently strong, such that weak instruments might be of concern. Given that including fixed effects does not improve the model fit and that the estimated coefficients remain unchanged, we proceed by estimating only the specification without fixed effects. This approach can also be justified on grounds of measurement aspects: If autocorrelated measurement error is present in the explanatory variables, including fixed effects reduces the variation of the signal, while increasing the variation of the measurement error. This exacerbates the potential bias through measurement error (Hauk and Wacziarg, 2009).

Table 3: First-Stage Results: Health and Schooling Instrumented

	Health Instrumented		Schooling Instrumented		Health and Schooling Instrumented			
Dependent variable	$\Delta(h_{i,t})$	$\Delta(h_{i,t})$	$\Delta(s_{i,t})$	$\Delta(s_{i,t})$	$\Delta(h_{i,t})$	$\Delta(s_{i,t})$	$\Delta(h_{i,t})$	$\Delta(s_{i,t})$
	(1)	(2)	(3)	(4)	(5.1)	(5.2)	(6.1)	(6.2)
$(\text{Child mortality})_{i,t-1}^{10-14}$	0.19***	0.22***	—	—	0.19***	-0.36	0.22***	-0.13
	(0.036)	(0.062)	—	—	(0.036)	(0.25)	(0.077)	(0.45)
$(\text{Schooling})_{i,t-1}^{55-59}$	—	—	-0.14***	-0.10***	-0.0035**	-0.13***	-0.00024	-0.10***
	—	—	(0.024)	(0.025)	(0.0015)	(0.025)	(0.0034)	(0.026)
R^2	0.27	0.19	0.23	0.17	0.27	0.24	0.19	0.17
First-stage F	27.2	12.8	34.4	15.7	16.8	23.4	6.7	7.9
Combined F	—	—	—	—	15.9		3.3	
Countries	116	116	116	116	116	116	116	116
Observations	613	613	613	613	613	613	613	613
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed Effects	—	✓	—	✓	—	—	✓	✓

Notes: The dependent variable is either the change in adult survival $\Delta(h_{i,t})$ or the change in years of secondary schooling $\Delta(s_{i,t})$. All specifications include lagged output per worker $ln(y_{i,t-1})$, the growth rate of capital per worker $\Delta ln(k_{i,t})$, changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$, lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. Even columns allow for country-specific growth trends by including fixed effects. First-stage *F* reports the Kleibergen-Paap *F*-statistic. Combined *F* in Columns (5) and (6) refers to the first-stage F-statistic of the combined model, in which health and schooling are jointly instrumented. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

The first three columns of Table 4 report the second-stage results. In the specifications in Columns (1) and (2), we instrument changes in health and education inputs separately, whereas we jointly instrument them in the specification in Column (3). Qualitatively, the estimated coefficients have the same sign as in the OLS specifications, again confirming theoretical predictions and results of the growth literature. Quantitatively, the coefficients of all variables take similar values across all three specifications and, moreover, are close to the results obtained from OLS. According to Column (3), the estimated coefficient of population health is 0.65, implying a return on health of approximately 1 percent. Correspondingly, an increase in the adult survival rate by 0.1 is associated with a 10-percent increase in labor productivity. The estimated coefficient of workforce education is slightly smaller than in the OLS estimation and takes a value of roughly 0.05. This implies a return of an additional year of secondary education on labor productivity of roughly 7.7 percent—a value that lies in the middle quintile of Mincerian returns reported in the literature (Psacharopoulos and Patrinos, 2018). Hence, the estimated returns on both population health and education fall in the range of plausible microeconomic estimates. A caveat of our IV specifications, however, is that the estimated coefficients for health and education inputs are insignificant. Hence, the model lacks the statistical power to determine whether health and education have a positive effect on labor productivity. While we take this concern seriously, we consider the evidence to indicate that the coefficient estimates are quantitatively very similar despite the substantial uncertainty in the data. Therefore, we view our IV results as another piece of evidence that micro- and macro-based estimates of the return on health are consistent with each other.

Table 4: Return of Health and Education to Productivity: IV Results

	Two-Stage Least Squares			Panel GMM		Panel GMM + IV	
	$\Delta(h_{i,t})$ Instrumented (1)	$\Delta(s_{i,t})$ Instrumented (2)	$\Delta(h_{i,t})$, $\Delta(s_{i,t})$ Instrumented (3)	Difference GMM (4)	System GMM (5)	Difference GMM (6)	System GMM (7)
$\ln(y_{i,t-1})$	-0.15***	-0.15***	-0.15***	-0.20***	-0.048	-0.20***	-0.086**
	(0.020)	(0.016)	(0.020)	(0.061)	(0.046)	(0.061)	(0.040)
$\Delta \ln(k_{i,t})$	0.35***	0.35***	0.35***	0.23	0.45 ***	0.24*	0.41***
	(0.062)	(0.064)	(0.065)	(0.15)	(0.078)	(0.14)	(0.070)
$\Delta(h_{i,t})$	0.74	0.59**	0.65	0.84**	0.78**	0.79*	0.63**
	(1.01)	(0.29)	(0.98)	(0.41)	(0.31)	(0.40)	(0.32)
$\Delta(s_{i,t})$	0.075**	0.050	0.049	0.11*	0.067**	0.12**	0.075**
	(0.030)	(0.084)	(0.083)	(0.061)	(0.031)	(0.059)	(0.030)
$\Delta(a_{i,t})$	0.0073	0.0076	0.0075	-0.0059	0.0089	-0.0055	0.0082
	(0.0084)	(0.0085)	(0.0084)	(0.011)	(0.0099)	(0.011)	(0.011)
$\Delta(a_{i,t}^2)$	-0.00057	-0.00056	-0.00056	0.00012	-0.00057	0.00011	-0.00056
	(0.00035)	(0.00035)	(0.00035)	(0.00051)	(0.00051)	(0.00053)	(0.00053)
R^2	0.38	0.38	0.38	—	—	—	—
First-stage F	27.2	34.4	15.9	—	—	—	—
AR(2) p–value	—	—	—	0.3	0.2	0.3	0.2
Hansen p–value	—	—	—	0.04	0.09	0.04	0.07
Diff.-in-Hansen p–value	—	—	—	—	0.77	—	0.62
Instruments	1	1	2	77	89	79	93
Countries	116	116	116	116	116	116	116
Observations	613	613	613	497	613	497	613
Controls	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The panel GMM specifications instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The difference GMM specification uses up to two lags of the endogenous variables, whereas the system GMM specification uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. Country-fixed effects in the GMM specifications are removed using forward-orthogonal deviations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). Columns (6) and (7) add the nondemeaned external instruments to the GMM specifications from Columns (4) and (5). First-stage *F* refers to the Kleibergen-Paap *F*-statistic. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

6.3 Panel GMM Results

As a further identification strategy, we use dynamic panel GMM estimators to address potential spurious correlations that arise mechanically through a link between the lagged dependent variable and the error term. The difference GMM estimator identifies the parameters of interest by instrumenting changes in the explanatory variables by their lagged values (Arellano and Bond, 1991). The system GMM estimator extends this model by a level equation, which exploits changes over time as additional instruments, to obtain efficient estimates (Arellano and Bover, 1995; Blundell and Bond, 1998).[12] In our specifications, we instrument not only the lagged dependent variable but also some of the other regressors by lagged values, because these inputs potentially correlate with productivity shocks over five-year intervals. Hence, this approach is more cautious than a parsimonious GMM specification that only instruments the lagged dependent variable. Specifically, we instrument lagged output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in average years of secondary schooling $\Delta(s_{i,t})$. By contrast, we treat changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}^2)$ as exogenous variables. This assumption is appropriate, because, on the one hand, the demographic structure is strongly predetermined and thus insensitive to contemporaneous shocks, and on the other hand, compulsory schooling laws require sufficient time to alter the schooling composition of a country.

An important aspect of both difference and system GMM is the choice of the instrument sets: While the estimators' efficiency increases with the number of instruments, the risk of using weak instruments also increases. A rule of thumb due to Roodman (2009) is not to use more instruments than there are cross-sectional units in the sample—here countries. Therefore, we restrict the instrument count by using only the two most recent lags of the endogenous variables in the difference GMM specification. For system GMM, which has a higher instrument count through the inclusion of the levels equation, we use only the most recent lag of the endogenous variables in the difference equation and the first difference of the endogenous variables in the level equation.

Columns (4) and (5) of Table 4 report the results for difference and system GMM. Both estimators eliminate fixed effects using forward-orthogonal deviations and thus also account for country-specific growth trends. The AR(2) test shows that there is no second-order autocorrelation in the nondemeaned error terms, such that it is possible to use the twice-lagged levels of the endogenous variables as instrument. Overall, the estimates accord with the OLS results in Table 1 and with the IV results in the first three columns of Table 4. The computed coefficients of all variables have a similar magnitude as the OLS and IV estimates. In particular, the reduced-form estimates of population health are significantly different from zero at the 5-percent level and take values of roughly 0.8. Hence, an increase in the adult survival rate by 0.1 is associated with an increase in labor productivity slightly above 10 percent. This value again lies in the range of plausible estimates of the microeconomic return on health. Similarly, the estimated coefficients of workforce education also take similar values as in the IV and OLS regressions, with difference GMM yielding a somewhat higher return on education than system GMM. Nonetheless, both values fall in the range of plausible microeconomic estimates reported in the literature. Due to the higher number of instruments, the GMM estimates are more precise than the IV estimates. We view the results

[12] See also Holtz-Eakin et al. (1988) for further details.

as a further piece of evidence that macro-based estimates of the economic return on health are compatible with micro-based estimates, for example those by Weil (2007). The Hansen J-test, however, indicates that the instrument set is not perfectly exogenous—in particular for the less precise difference GMM estimator—such that these estimates should not be interpreted as causal effects. We further address this issue in the robustness section, where we test alternative GMM specifications to gauge how sensitively the results react to different instrument sets.

Finally, Columns (6) and (7) contain specifications in which we combine the nondemeaned external instruments for health and education inputs with the GMM instruments from Columns (4) and (5). The results conform quantitatively and qualitatively to the GMM estimates without external instruments. As before, the estimated coefficients fall in the same domain as the OLS and IV results, with the return on population health estimated by system GMM being almost identical.

6.4 Robustness

This section shows the robustness of the results for alternative measures of income, education, and health and alternative panel lengths and discusses the sensitivity of the GMM results with respect to different specifications. The Appendix presents the corresponding tables.

Total Years of Schooling. As a first robustness check, Table A2 presents the estimation results of an empirical model that proxies education by average years of total schooling instead of average years of secondary schooling. Qualitatively and quantitatively, the results conform closely to our main results. In particular, the estimated economic return on health is quantitatively almost identical to the main results and lies in the range of plausible microeconomic estimates. Accordingly, micro-based and macro-based estimates of the economic return to health are consistent with each other.

At the same time, the estimated return on an additional year of schooling is statistically insignificant in some specifications; however, with values of 5 to 7 percent, it still lies at the lower end of the range of plausible values of microeconomic estimates in quantitative terms. The variation in this regressor may be less informative than the variation in years of secondary education, because primary education and tertiary education vary less over the observation period than secondary years of schooling.

GDP per Capita. As a second robustness check, Table A3 contains the estimation results for the growth rate of per capita GDP as the dependent variable instead of the growth rate of per worker GDP. In these specifications, we additionally control for the log difference in the workforce to population ratio $\Delta ln\left(L_{i,t}/N_{i,t}\right)$, thereby following the derivation of our model in per capita terms as in Equation (3). Overall, the estimated economic return to health is quantitatively very similar to the main results.

Life Expectancy. As a third robustness check, Table A4 reports results for empirical specifications, in which we proxy health inputs by life expectancy at birth instead of the adult survival rate. The magnitude of the corresponding return to health cannot be directly compared with the other specifications. Nonetheless, these results are informative, because life expectancy by construction depends on child and adult mortality, such that it provides an alternative proxy for the health of the workforce. The estimates qualitatively confirm the empirical patterns observed in the main results. Given the conceptual relation between life expectancy and adult survival rates, we view this stability of the qualitative results as a reassuring piece of evidence of robustness, corroborating our main findings.

10-Year Panels. As a fourth robustness check, Table A5 shows results for 10-year panel intervals instead of five-year intervals. As this approach requires country observations over longer time periods, the estimation sample shrinks to 111 countries for OLS, IV, and system GMM regressions and to 90 countries for difference GMM specifications. The results qualitatively and quantitatively confirm the main results. In particular, the implied economic returns to health are very similar across specifications and again conform to Weil's (2007) point estimate and the range of plausible microeconomic estimates. Compared with the main results, the coefficients are, however, less precisely estimated and statistically insignificant in some specifications.

Robustness GMM. As a final robustness check, Tables A6 and A7 report results for alternative GMM specifications. The rationale for these specification tests is that the empirical results in GMM models may sensitively react to changes in the instrument set. Specifically, we estimate models that differ from our baseline GMM specification in the following dimensions: We estimate models with fewer lag periods, more lag periods, models that additionally use the external instruments, models that use the external instruments instead of the GMM instruments, models that collapse the instrument set, models that eliminate fixed effects using first differences instead of forward-orthogonalized deviations, and models that instrument only lagged output per worker. Overall, the corresponding results are qualitatively and quantitatively similar to our main findings. In particular, the estimated return on health again falls in the range of plausible microeconomic values. Hence, this evidence again indicates that micro-based and macro-based estimates of the economic return on health are consistent with each other.

7 Conclusion

The growth literature has used micro-based and macro-based approaches to assess the effects of health on economic growth. Micro-based approaches aggregate the return on health obtained from Mincerian wage regressions to derive the macroeconomic return on health, whereas macro-based approaches estimate generalized production functions decomposing output into its components. Micro-based approaches tend to find substantially smaller effects of health on economic performance than the macro-based approaches, thereby raising a micro-macro puzzle of the economic return on health. We believe our study is the first to show that the macroeconomic estimates of the effect of health on output are compatible with the microeconomic estimates of the effect of health on wages. According to our estimates, an increase in adult survival rates of 0.1, or 10 percent, increases labor productivity by about 9 to 10 percent, which is somewhat higher than, but still consistent with, Weil's (2007) calibrated value of around 6.7 percent. Apart from the fact that Weil's (2007) reported estimate relies on the two most conservative studies at the microeconomic level, the slightly higher point estimate from the macro-based approach might also be due to spillover effects that the micro-based approach omits by design. Given that we find no evidence of substantial externalities, however, our results suggest that calibration based on microeconomic data can serve as a reasonable means to estimate the macroeconomic impact of health changes. Thus, our results provide a macro-based justification for using a micro-based approach to estimate the direct economic benefits of specific health interventions.

Overall our results indicate that health plays a role in explaining cross-country differences in the level of income per worker. As far as policy implications are concerned, public health measures might be an important lever for fostering economic development. Potential policies along these lines include vaccination programs, antibiotic distribution programs, and micronutrient supplementation schemes, which lead to large improvements in health outcomes for relatively low expenditures (World Bank, 1993; Commission on Macroeconomics and Health, 2001; Field et al., 2009; Luca et al., 2014).

References

Acemoglu, D. and Johnson, S. (2007). Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy*, 115(6):925–985.

Aghion, P., Howitt, P., and Murtin, F. (2011). The relationship between health and growth: When Lucas meets Nelson-Phelps. *Review of Economics and Institutions*, 2(1):1–24.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106(2):407–443.

Barro, R. J. (1997). *Determinants of economic growth: A cross-country empirical study*. MIT Press, Cambridge, MA.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104:184–198.

Barro, R. J. and Sala-i-Martin, X. (2004). *Economic growth*. MIT Press, Cambridge, MA.

Baumol, W. J. (1986). Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, 76(5):1072–1085.

Bhargava, A., Jamison, D. T., Lau, L. J., and Murray, C. J. (2001). Modeling the effects of health on economic growth. *Journal of Health Economics*, 20(3):423–440.

Bils, M. and Klenow, P. J. (2000). Does schooling cause growth? *American Economic Review*, 90(5):1160–1183.

Bleakley, H. (2007). Disease and development: Evidence from hookworm eradication in the American South. *Quarterly Journal of Economics*, 122(1):73–117.

Bleakley, H. (2010). Health, human capital, and development. *Annual Review of Economics*, 2:283–310.

Bleakley, H. and Lange, F. (2009). Chronic disease burden and the interaction of education, fertility and growth. *Review of Economics and Statistics*, 91(1):52–65.

Bloom, D. E. and Canning, D. (2000). The health and wealth of nations. *Science*, 287(5456):1207–1209.

Bloom, D. E., Canning, D., and Fink, G. (2014a). Disease and development revisited. *Journal of Political Economy*, 122(6):1355–1366.

Bloom, D. E., Canning, D., and Graham, B. (2003a). Longevity and life-cycle savings. *Scandinavian Journal of Economics*, 105(3):319–338.

Bloom, D. E., Canning, D., Mansfield, R. K., and Moore, M. (2007). Demographic change, social security systems, and savings. *Journal of Monetary Economics*, 54(1):92–114.

Bloom, D. E., Canning, D., and Moore, M. (2014b). Optimal retirement with increasing longevity. *Scandinavian Journal of Economics*, 116(3):838–858.

Bloom, D. E., Canning, D., and Sevilla, J. (2003b). *The demographic dividend: A new perspective on the economic consequences of population change*. Rand Corporation, Santa Monica, CA.

Bloom, D. E., Canning, D., and Sevilla, J. (2004). The effect of health on economic growth: A production function approach. *World Development*, 32(1):1–13.

Bloom, D. E., Kuhn, M., and Prettner, K. (2015). The contribution of female health to economic development. *NBER Working Paper 21411*, Cambridge, MA.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3):321–340.

Brückner, M. (2013). On the simultaneity problem in the aid and growth debate. *Journal of Applied Econometrics*, 28(1):126–150.

Case A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics,* 24(2): 365–389.

Caselli, F., Esquivel, G., and Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*, 1(3):363–389.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies*, 32(3):233–240.

Cervellati, M. and Sunde, U. (2005). Human capital formation, life expectancy, and the process of development. *American Economic Review*, 95(5):1653–1672.

Cervellati, M. and Sunde, U. (2011). Life expectancy and economic growth: The role of the demographic transition. *Journal of Economic Growth*, 16(2):99–133.

Cervellati, M. and Sunde, U. (2013). Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. *Econometrica*, 81(5):2055–2086.

Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: Investing in health for economic development*. World Health Organization, Geneva.

Cuaresma, J. C., Lutz, W., and Sanderson, W. (2014). Is the demographic dividend an education dividend? *Demography*, 51(1):299–315.

Currie, J., Stabile, M., Manivong, P., and Roos, L.L. (2010). Child health and young adult outcomes. *Journal of Human Resources*, 45(3): 517–548.

De la Fuente, A. and Domenech, R. (2001). Schooling data, technological diffusion, and the neoclassical model. *American Economic Review, Papers & Proceedings*, 91(2):323–327.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review*, 55(5):1126–1150.

Doppelhofer, G., Miller, R. I., and Sala-i-Martin, X. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review*, 94(4):813–835.

Dowrick, S. and Rogers, M. (2002). Classical and technological convergence: Beyond the Solow-Swan growth model. *Oxford Economic Papers*, 54(3):369–385.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. (2005). Growth econometrics. In *Handbook of Economic Growth*, volume 1A, chapter 8, pages 555–677. North Holland, Amsterdam.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics*, 112(4):1203–1250.

Eberhardt, M. and Teal, F. (2011). Econometrics for grumblers: A new look at the literature on cross-country growth empirics. *Journal of Economic Surveys*, 25(1):109–155.

Fagerberg, J. (1994). Technology and international differences in growth rates. *Journal of Economic Literature*, 32(3):1147–1175.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the Penn world table. *American Economic Review*, 105(10):3150–3182.

Field, E., Robles, O., and Torero, M. (2009). Iodine deficiency and schooling attainment in Tanzania. *American Economic Journal: Applied Economics*, 1(4):140–169.

Fogel, R. W. (1994). Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy. *American Economic Review*, 84(3):369–395.

Fogel, R. W. (1997). New findings on secular trends in nutrition and mortality: Some implications for population theory. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*, volume 1A, chapter 9, pages 433–481. North Holland, Amsterdam.

Gallup, J. L., Sachs, J. D., and Mellinger, A. D. (1999). Geography and economic development. *International Regional Science Review*, 22(2):179–232.

Galor, O. (2005). From stagnation to growth: Unified growth theory. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1A, chapter 4, pages 171–293. North Holland, Amsterdam.

Galor, O. (2011). *Unified growth theory*. Princeton University Press, Princeton, NJ.

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828.

Griliches, Z. and Mairesse, J. (1998). Production functions: The search for identification. In *Econometrics and Economic Theory in the 20th Century*. Cambridge University Press, Cambridge.

Gwartney, J., Lawson, R., and Hall, J. (2017). *Economic freedom of the world: 2017 annual report*. Fraser Institute, Vancouver, BC.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1):83–116.

Hansen, C. W. and Lønstrup, L. (2015). The rise in life expectancy and economic growth in the 20th century. *Economic Journal*, 125(584):838–852.

Hauk, W. R. and Wacziarg, R. (2009). A Monte Carlo study of growth regressions. *Journal of Economic Growth*, 14(2):103–147.

Heckman, J. J. and Klenow, P. J. (1997). *Human capital policy*. University of Chicago Press, Chicago, IL.

Holtz-Eakin, D., Newey, W., and Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6):1371–1395 .

Islam, N. (1995). Growth empirics: A panel data approach. *Quarterly Journal of Economics*, 110(4):1127–1170.

Judson, R. A. and Owen, A. L. (1999). Estimating dynamic panel data models: A guide for macroeconomists. *Economics Letters*, 65(1):9–15.

Kotschy, R. and Sunde, U. (2018). Can education compensate the effect of population ageing on macroeconomic performance? *Economic Policy*, 33(96):587–634.

Krueger, A. B. and Lindahl, M. (2001). Education for growth: Why and for whom? *Journal of Economic Literature*, 39(4):1101–1136.

Levine, R. and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82(4):942–963.

Lorentzen, P., McMillan, J., and Wacziarg, R. (2008). Death and development. *Journal of Economic Growth*, 13(2):81–124.

Luca, D. L., Iversen, J., Lubet, A. S., Mitgang, E., Onarheim, K. H., Prettner, K., and Bloom, D. E. (2014). Benefits and costs of the women's health targets for the post-2015 development agenda. *Copenhagen Consensus Center Working Paper*.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107(2):407–437.

Miguel E. and Kremer M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72(1):159–217.

Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review: Papers & Proceedings*, 56(1):69–75.

Prettner, K. (2013). Population aging and endogenous economic growth. *Journal of Population Economics*, 26(2):811–834.

Pritchett, L. H. (2001). Where has all the education gone? *World Bank Economic Review*, 15(3):367–391.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2):111–134.

Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics*, 26(5):1–14.

Reynal-Querol, M. and Montalvo, J. G. (2005). Ethnic polarization, potential conflict, and civil war. *American Economic Review*, 95(3):796–816.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):71–102.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1): 135–158.

Sachs, J. D. and Warner, A. M. (1997). Sources of slow growth in African economies. *Journal of African Economies*, 6(3):335–376.

Sachs, J. D., Warner, A., Aslund, A., and Fischer, S. (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity*, 1995(1):1–118.

Sala-i-Martin, X. (1997). I just ran two million regressions. *American Economic Review: Papers & Proceedings*, 87(2):178–183.

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review*, 92(2):349–353.

Shastry, G. K. and Weil, D. N. (2003). How much of cross-country income variation is explained by health? *Journal of the European Economic Association*, 1(2–3):387–396.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94.

Solt, F. (2016a). The standardized world income inequality database. *Social Science Quarterly*, 97(5):1267–1281.

Solt, F. (2016b). The standardized world income inequality database. *Social Science Quarterly*, 97(5):1267–1281. SWIID Version 6.1, October 2017.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817.

Strulik, H., Prettner, K., and Prskawetz, A. (2013). The past and future of knowledge-based growth. *Journal of Economic Growth*, 18(4):411–437.

Sunde, U. and Vischer, T. (2015). Human capital and growth: Specification matters. *Economica*, 82(326):368–390.

UNESCO. (1963–1997). *Statistical yearbooks*. United Nations Educational, Scientific and Cultural Organization, Paris.

UNESCO. (2017). *Data for sustained development goals*. United Nations Educational, Scientific and Cultural Organization, Institute for Statistics, Paris.

United Nations. (2017). *World population prospects: The 2017 revision*. United Nations, Department of Economic and Social Affairs, Population Division, New York, NY.

Weil, D. N. (2007). Accounting for the effect of health on economic growth. *Quarterly Journal of Economics*, 122(3):1265–1306.

Weil, D. N. (2014). Health and economic growth. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2, chapter 3, pages 623–682. Elsevier, Amsterdam.

Weil, D. N. (2017). Health improvement and income growth in the long run. In Cervellati, M. and Sunde, U., editors, *Demographic Change and Long-Run Development*, pages 43–68. MIT Press, Cambridge, MA.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1):25–51.

World Bank. (1993). *World development report 1993: Investing in health*. World Bank, Washington, DC.

World Bank. (2017). *World development indicators 1960–2017*. World Bank, Washington, DC.

Appendix

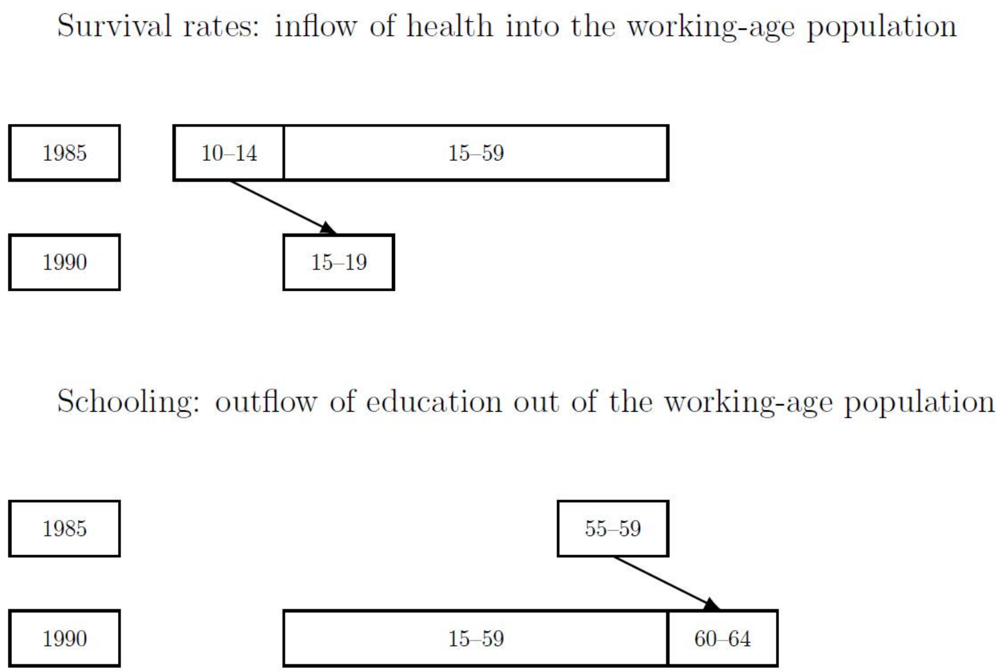

Figure A1: Demographic Dynamics as Instrumental Variables

Source: Illustraton adjusted from Kotschy and Sunde (2018, Online Appendix p. 4).

Table A1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Obs.
		GDP and physical capital			
Log output per worker	9.82	1.07	7.31	12.01	613
Growth of output per worker	0.10	0.21	-0.91	1.27	613
Log output per capita	8.83	1.16	6.21	11.71	613
Growth of output per capita	0.12	0.22	-0.79	1.23	613
Log capital per worker	11.10	1.23	7.27	13.08	613
Growth of capital per worker	0.10	0.16	-0.39	0.90	613
		Health and schooling			
Adult survival rate	0.79	0.11	0.28	0.94	613
Change in adult survival rate	0.01	0.03	-0.18	0.14	613
Life expectancy at birth	66.52	10.16	34.20	82.66	613
Change in life expectancy at birth	1.42	1.64	-7.31	9.47	613
Years of secondary schooling	2.61	1.56	0.06	7.96	613
Change in years of secondary schooling	0.27	0.25	-0.56	1.67	613
Years of total schooling	7.58	3.07	0.47	13.22	613
Change in years of total schooling	0.53	0.37	-0.75	2.23	613
		Instruments for health and schooling			
(Child mortality rate)$_{t-1}^{10-14}$	0.10	0.09	0.01	0.47	613
(Years of secondary schooling)$_{t-1}^{55-59}$	1.47	1.46	0.005	7.15	613
(Years of total schooling)$_{t-1}^{55-59}$	4.98	3.39	0.01	13.30	613
		Experience, inequality, and controls			
Experience	12.03	7.40	0.50	29.70	613
Change in experience	0.60	1.57	-11.40	5.60	613
Log workforce to population ratio	-0.51	0.11	-0.76	-0.15	613
Change in workforce to population ratio	0.02	0.02	-0.04	0.13	613
Income inequality	0.39	0.09	0.21	0.59	508
Change in income inequality	0.001	0.02	-0.07	0.07	461
Size of agricultural sector (% of GDP)	15.02	13.51	0.12	72.03	613
Percent land area in tropics	0.46	0.47	0	1	613
Institutional Quality	6.00	1.33	2.28	8.65	613

Table A2: Robustness: Total Years of Schooling

	OLS Baseline (1)	OLS with Fixed Effects (2)	$\Delta(h_{i,t})$, $\Delta(s_{i,t})$ Instrumented (3)	Difference GMM (4)	System GMM (5)	Difference GMM + IV (6)	System GMM + IV (7)
$\ln(y_{i,t-1})$	-0.15***	-0.33***	-0.15***	-0.20***	-0.056	-0.21***	-0.078**
	(0.017)	(0.052)	(0.020)	(0.054)	(0.035)	(0.053)	(0.032)
$\Delta\ln(k_{i,t})$	0.35***	0.27**	0.37***	0.26	0.46***	0.24	0.43***
	(0.063)	(0.11)	(0.069)	(0.17)	(0.080)	(0.17)	(0.071)
$\Delta(h_{i,t})$	0.63**	0.66**	0.44	0.87**	0.78**	0.81*	0.65**
	(0.30)	(0.27)	(1.17)	(0.39)	(0.30)	(0.41)	(0.30)
$\Delta(s_{i,t})$	0.038	0.035	-0.036	0.043	0.040*	0.041	0.041**
	(0.023)	(0.029)	(0.083)	(0.035)	(0.021)	(0.033)	(0.021)
$\Delta(a_{i,t})$	0.0061	-0.010	0.0089	-0.0042	0.0075	-0.0051	0.0073
	(0.0088)	(0.0098)	(0.0083)	(0.011)	(0.011)	(0.011)	(0.011)
$\Delta(a^2_{i,t})$	-0.00051	0.00028	-0.00057*	0.000099	-0.00054	0.00012	-0.00057
	(0.00036)	(0.00037)	(0.00033)	(0.00051)	(0.00054)	(0.00048)	(0.00053)
R^2	0.38	0.38	0.36	—	—	—	—
First-stage F	—	—	8.6	—	—	—	—
AR(2) p–value	—	—	—	0.3	0.2	0.3	0.2
Hansen p–value	—	—	—	0.02	0.2	0.03	0.2
Diff.-in-Hansen p–value	—	—	—	—	0.91	—	0.84
Instruments	—	—	2	77	89	79	93
Countries	116	116	116	116	116	116	116
Observations	613	613	613	497	613	497	613
Controls	✓	✓	✓	✓	✓	✓	✓
Fixed Effects	—	✓	—	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of total schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The panel GMM specifications instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of total schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The difference GMM specification uses up to two lags of the endogenous variables, whereas the system GMM specification uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. Country-fixed effects in the GMM specifications are removed using forward-orthogonal deviations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). Columns (6) and (7) add the nondemeaned external instruments to the GMM specifications from Columns (4) and (5). First-stage *F* refers to the Kleibergen-Paap *F*-statistic. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A3: Robustness: GDP per Capita

	OLS Baseline (1)	OLS with Fixed Effects (2)	$\Delta(h_{i,t})$, $\Delta(s_{i,t})$ Instrumented (3)	Difference GMM (4)	System GMM (5)	Difference GMM + IV (6)	System GMM + IV (7)
$\ln(y_{i,t-1})$	-0.15***	-0.35***	-0.15***	-0.22***	-0.033	-0.22***	-0.075***
	(0.017)	(0.055)	(0.019)	(0.067)	(0.038)	(0.068)	(0.028)
$\Delta\ln(k_{i,t})$	0.16*	0.022	0.15*	-0.032	0.24***	-0.015	0.22***
	(0.079)	(0.12)	(0.077)	(0.15)	(0.082)	(0.14)	(0.081)
$\Delta(h_{i,t})$	0.67**	0.54*	0.52	0.48	0.86***	0.45	0.72**
	(0.30)	(0.28)	(0.91)	(0.42)	(0.33)	(0.41)	(0.34)
$\Delta(s_{i,t})$	0.074**	0.10**	0.099	0.13**	0.071**	0.14**	0.068*
	(0.033)	(0.044)	(0.077)	(0.062)	(0.035)	(0.061)	(0.036)
$\Delta(a_{i,t})$	0.0013	-0.0099	0.0017	-0.0048	0.0029	-0.0040	0.0013
	(0.010)	(0.0100)	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)
$\Delta(a^2_{i,t})$	-0.00034	0.00020	-0.00036	-0.0000062	-0.00045	-0.000042	-0.00033
	(0.00037)	(0.00037)	(0.00037)	(0.00053)	(0.00051)	(0.00055)	(0.00047)
R^2	0.36	0.39	0.36	—	—	—	—
First-stage F	—	—	13.8	—	—	—	—
AR(2) p–value	—	—	—	0.2	0.2	0.2	0.2
Hansen p–value	—	—	—	0.06	0.2	0.07	0.1
Diff.-in-Hansen p–value	—	—	—	—	0.92	—	0.86
Instruments	—	—	2	78	90	80	94
Countries	116	116	116	116	116	116	116
Observations	613	613	613	497	613	497	613
Controls	✓	✓	✓	✓	✓	✓	✓
Fixed Effects	—	✓	—	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per capita $\Delta ln(\tilde{y}_{i,t})$. All specifications include the log difference in the workforce to population ratio $\Delta ln(L_{i,t}/N_{i,t})$, lagged years of secondary schooling $s_{i,t-1}$, and a full set of time effects. The panel GMM specifications instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The difference GMM specification uses up to two lags of the endogenous variables, whereas the system GMM specification uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. Country-fixed effects in the GMM specifications are removed using forward-orthogonal deviations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). Columns (6) and (7) add the nondemeaned external instruments to the GMM specifications from Columns (4) and (5). First-stage *F* reports the Kleibergen-Paap *F*-statistic. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A4: Robustness: Life Expectancy

	OLS Baseline (1)	OLS with Fixed Effects (2)	$\Delta(h_{i,t})$, $\Delta(s_{i,t})$ Instrumented (3)	Difference GMM (4)	System GMM (5)	Difference GMM + IV (6)	System GMM + IV (7)
$\ln(y_{i,t-1})$	-0.15***	-0.34***	-0.15***	-0.23***	-0.062	-0.22***	-0.091**
	(0.017)	(0.052)	(0.019)	(0.058)	(0.040)	(0.058)	(0.036)
$\Delta\ln(k_{i,t})$	0.34***	0.29***	0.35***	0.23	0.44***	0.23*	0.41***
	(0.063)	(0.10)	(0.065)	(0.15)	(0.083)	(0.14)	(0.076)
$\Delta(h_{i,t})$	0.011*	0.013**	0.0079	0.016*	0.013**	0.015*	0.0100
	(0.0058)	(0.0056)	(0.012)	(0.0087)	(0.0060)	(0.0085)	(0.0063)
$\Delta(s_{i,t})$	0.076**	0.093**	0.054	0.12**	0.066**	0.12**	0.074***
	(0.031)	(0.046)	(0.083)	(0.054)	(0.030)	(0.057)	(0.028)
$\Delta(a_{i,t})$	0.0079	-0.0091	0.0081	-0.0028	0.0068	-0.0033	0.0058
	(0.0087)	(0.0094)	(0.0084)	(0.011)	(0.0096)	(0.011)	(0.0099)
$\Delta(a_{i,t}^2)$	-0.00058	0.00021	-0.00057	0.000059	-0.00048	0.000060	-0.00043
	(0.00036)	(0.00036)	(0.00035)	(0.00049)	(0.00052)	(0.00051)	(0.00052)
R^2	0.38	0.38	0.38	—	—	—	—
First-stage F	—	—	16.3	—	—	—	—
AR(2) p–value	—	—	—	0.3	0.2	0.3	0.2
Hansen p–value	—	—	—	0.05	0.10	0.04	0.08
Diff.-in-Hansen p–value	—	—	—	—	0.88	—	0.75
Instruments	—	—	2	77	89	79	93
Countries	116	116	116	116	116	116	116
Observations	613	613	613	497	613	497	613
Controls	✓	✓	✓	✓	✓	✓	✓
Fixed Effects	—	✓	—	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The panel GMM specifications instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The difference GMM specification uses up to two lags of the endogenous variables, whereas the system GMM specification uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. Country-fixed effects in the GMM specifications are removed using forward-orthogonal deviations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). Columns (6) and (7) add the nondemeaned external instruments to the GMM specifications from Columns (4) and (5). First-stage *F* reports the Kleibergen-Paap *F*-statistic. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A5: Robustness: 10-Year Panel Intervals

	OLS Baseline (1)	OLS with Fixed Effects (2)	$\Delta(h_{i,t})$, $\Delta(s_{i,t})$ Instrumented (3)	Difference GMM (4)	System GMM (5)	Difference GMM + IV (6)	System GMM + IV (7)
$\ln(y_{i,t-1})$	-0.25***	-0.54***	-0.24***	-0.30	-0.13*	-0.28	-0.18***
	(0.026)	(0.10)	(0.028)	(0.23)	(0.067)	(0.22)	(0.066)
$\Delta \ln(k_{i,t})$	0.34***	0.26**	0.34***	0.047	0.41***	0.077	0.40***
	(0.064)	(0.11)	(0.063)	(0.23)	(0.091)	(0.24)	(0.091)
$\Delta(h_{i,t})$	0.49	0.68*	0.70	0.67	0.73**	0.68	0.69**
	(0.30)	(0.37)	(1.11)	(0.61)	(0.32)	(0.57)	(0.35)
$\Delta(s_{i,t})$	0.072**	0.082	0.025	0.081	0.044	0.094	0.040
	(0.036)	(0.063)	(0.095)	(0.12)	(0.044)	(0.12)	(0.044)
$\Delta(a_{i,t})$	0.017	-0.0056	0.016	-0.0010	0.023	-0.00089	0.018
	(0.010)	(0.013)	(0.010)	(0.022)	(0.017)	(0.021)	(0.018)
$\Delta(a^2_{i,t})$	-0.0011**	0.00016	-0.0010**	-0.000090	-0.0017*	-0.000067	-0.0015
	(0.00051)	(0.00060)	(0.00049)	(0.0012)	(0.00099)	(0.0011)	(0.00099)
R^2	0.52	0.51	0.52	—	—	—	—
First-stage F	—	—	12.7	—	—	—	—
AR(2) p–value	—	—	—	0.9	0.6	1.0	0.6
Hansen p–value	—	—	—	0.04	0.008	0.06	0.007
Diff.-in-Hansen p–value	—	—	—	—	0.06	—	0.05
Instruments	—	—	2	32	44	34	48
Countries	111	111	111	90	111	90	111
Observations	297	297	297	186	297	186	297
Controls	✓	✓	✓	✓	✓	✓	✓
Fixed Effects	—	✓	—	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The panel GMM specifications instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The difference GMM specification uses up to two lags of the endogenous variables, whereas the system GMM specification uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. Country-fixed effects in the GMM specifications are removed using forward-orthogonal deviations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). Columns (6) and (7) add the nondemeaned external instruments to the GMM specifications from Columns (4) and (5). First-stage *F* reports the Kleibergen-Paap *F*-statistic. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A6: Robustness: Alternative Difference GMM Specifications

	OLS with Fixed Effects (1)	Baseline GMM (2)	Fewer Lags (3)	More Lags (4)	GMM & External IVs (5)	Only External IVs (6)	Collapsed Instrument Set (7)	First Differences (8)	Only GDP Instrumented (9)
$\ln(y_{i,t-1})$	-0.34***	-0.20***	-0.18**	-0.26***	-0.20***	-0.27***	-0.23***	-0.29***	-0.35***
	(0.052)	(0.061)	(0.079)	(0.054)	(0.061)	(0.098)	(0.069)	(0.086)	(0.086)
$\Delta\ln(k_{i,t})$	0.29***	0.23	0.55	0.20	0.24*	0.0039	0.14	0.38*	0.32***
	(0.10)	(0.15)	(0.50)	(0.13)	(0.14)	(0.20)	(0.19)	(0.21)	(0.10)
$\Delta(h_{i,t})$	0.64**	0.84**	0.96	0.80**	0.79*	2.97***	0.60	0.29	0.65**
	(0.27)	(0.41)	(0.79)	(0.34)	(0.40)	(0.98)	(0.53)	(0.36)	(0.33)
$\Delta(s_{i,t})$	0.091**	0.11*	-0.013	0.11**	0.12**	-0.17	0.11	0.100	0.083*
	(0.046)	(0.061)	(0.15)	(0.050)	(0.059)	(0.19)	(0.069)	(0.15)	(0.045)
$\Delta(a_{i,t})$	-0.0100	-0.0059	0.0057	-0.0096	-0.0055	-0.010	-0.0055	-0.0050	-0.0082
	(0.0095)	(0.011)	(0.012)	(0.011)	(0.011)	(0.013)	(0.014)	(0.0096)	(0.011)
$\Delta(a^2_{i,t})$	0.00024	0.00012	-0.00013	0.00021	0.00011	0.00023	0.000042	0.00017	0.00024
	(0.00036)	(0.00051)	(0.00053)	(0.00046)	(0.00053)	(0.00054)	(0.00060)	(0.00043)	(0.00049)
AR(2) p–value	—	0.3	0.5	0.3	0.3	0.7	0.3	0.3	0.3
Hansen p–value	—	0.04	0.008	0.4	0.04	0.02	0.001	0.03	0.02
Instruments	—	77	45	109	79	47	59	76	76
Countries	116	116	116	116	116	116	116	116	116
Observations	613	497	497	497	497	497	497	490	497
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The difference GMM specifications in Columns (2) to (9) instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The baseline GMM specification in Column (2) uses up to two lags of the endogenous variables—see Section 6.3. The specification in Column (3) uses one lag period less and the specification in Column (4) uses one lag period more compared with the baseline specification. Column (5) adds the nondemeaned external instruments to the baseline GMM specification. The specification in Column (6) uses GMM instruments for all endogenous variables except health and schooling, which are instrumented only by the external instruments. For the specification in Column (7), the lag count is limited by collapsing the instrument set. Country-fixed effects in the GMM specifications are removed using first differences in Column (8) and forward-orthogonal deviations in all other columns. The specification in Column (9) only instruments lagged output per worker. The number of observations refers to the transformed data. The AR(2) p-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A7: Robustness: Alternative System GMM Specifications

	OLS with Fixed Effects (1)	Baseline GMM (2)	More Lags (3)	GMM & External IVs (4)	Only External IVs (5)	Collapsed Instrument Set (6)	First Differences (7)	Only GDP Instrumented (8)
$\ln(y_{i,t-1})$	-0.34***	-0.048	-0.13***	-0.086**	-0.073	-0.18***	-0.053	-0.20***
	(0.052)	(0.046)	(0.039)	(0.040)	(0.053)	(0.048)	(0.053)	(0.043)
$\Delta \ln(k_{i,t})$	0.29***	0.45***	0.37***	0.41***	0.40***	0.31***	0.42***	0.33***
	(0.10)	(0.078)	(0.075)	(0.070)	(0.095)	(0.098)	(0.080)	(0.082)
$\Delta(h_{i,t})$	0.64**	0.78**	0.50	0.63**	1.72**	0.43	0.66**	0.62**
	(0.27)	(0.31)	(0.74)	(0.32)	(0.70)	(0.44)	(0.32)	(0.30)
$\Delta(s_{i,t})$	0.091**	0.067**	0.073	0.075**	0.052	0.075*	0.057*	0.077**
	(0.046)	(0.031)	(0.049)	(0.030)	(0.12)	(0.045)	(0.032)	(0.034)
$\Delta(a_{i,t})$	-0.0100	0.0089	0.0047	0.0082	0.0050	0.00081	0.0042	0.0025
	(0.0095)	(0.0099)	(0.013)	(0.011)	(0.011)	(0.012)	(0.0093)	(0.0094)
$\Delta(a_{i,t}^2)$	0.00024	-0.00057	-0.00055	-0.00056	-0.00046	-0.00024	-0.00019	-0.00039
	(0.00036)	(0.00051)	(0.00070)	(0.00053)	(0.00056)	(0.00058)	(0.00039)	(0.00047)
AR(2) p–value	—	0.2	0.2	0.2	0.5	0.2	0.2	0.2
Hansen p–value	—	0.09	0.4	0.07	0.01	0.001	0.2	0.08
Diff.-in-Hansen p–value	—	0.77	1.00	0.62	0.52	0.17	0.9	0.4
Instruments	—	89	121	93	59	71	89	93
Countries	116	116	116	116	116	116	116	116
Observations	613	613	613	613	613	613	613	613
Controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta ln(y_{i,t})$. All specifications include lagged years of secondary schooling $s_{i,t-1}$, controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, a full set of regional dummies, and a full set of time effects. The system GMM specifications in Columns (2) to (8) instrument lagged log output per worker $ln(y_{i,t-1})$, the growth rate of physical capital per worker $\Delta ln(k_{i,t})$, the change in population health $\Delta(h_{i,t})$, and the change in years of secondary schooling $\Delta(s_{i,t})$, whereas the changes in experience $\Delta(a_{i,t})$ and experience squared $\Delta(a_{i,t}{}^2)$ are treated as exogenous variables. The baseline GMM specification in Column (2) uses the first lag of the endogenous variables in the differences equation and the first difference of the endogenous variables in the levels equation—see Section 6.3. The specification in Column (3) uses one lag period more compared with the baseline specification. Column (4) adds the nondemeaned external instruments to the baseline GMM specification. The specification in Column (5) uses GMM instruments for all endogenous variables except health and schooling, which are instrumented only by the external instruments. For the specification in Column (6), the lag count is limited by collapsing the instrument set. Country-fixed effects in the GMM specifications are removed using first differences in Column (7) and forward-orthogonal deviations in all other columns. The specification in Column (8) only instruments lagged output per worker. The number of observations refers to the untransformed data. The AR(2) *p*-value refers to the null hypothesis that there is no second-order autocorrelation in the nondemeaned error terms. Standard errors in GMM are computed with the two-step procedure and corrected with respect to finite sample size (Windmeijer, 2005). The standard errors in all specifications are clustered on the country level. Asterisks indicate significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table A8: Coding Description

This table describes coding choices for countries in which compulsory schooling laws differ by schooling type or target group and countries that experienced longer spells of turbulence and civil war. This list contains all countries for which information on compulsory schooling was available and thus even those that do not enter the estimation sample.
Albania: United Nations Educational, Scientific and Cultural Organization (UNESCO) yearbooks report four plus an additional three years of compulsory schooling for 1963 and 1964. From 1965 to 1967, four plus an additional four years of compulsory schooling are reported. We code these as seven and eight years of schooling because "[f]our years' schooling is compulsory for all children; a second period of three (four) years is compulsory for children in towns and villages where a seven-grade (eight-grade) school is available" (UNESCO, 1963–1968).
Andorra: Andorra's educational system is split into French and Spanish schools. However, because both schooling systems differ in terms of compulsory schooling, we follow UNESCO's convention and code values as missing until 1977. Afterward, both schools require a minimum of 10 years of schooling so that we code a value of 10.
Angola: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as is the case for Portugal.
Argentina: In 1972, compulsory schooling takes a value of eights years while before and afterward compulsory schooling is consistently reported with seven years. Because the structure of the educational system did not change in 1972, we code a value of seven years.
Australia: For 1963 and 1964, UNESCO yearbooks report eight to 10 years of compulsory schooling, varying by state. We take the figure of New South Wales, the most densely populated state, and thus code nine years. From 1968 onward, the yearbooks report values between nine and 11 years, varying by state or whether kindergarten counts toward primary education. We code 10 years of compulsory schooling. This figure is consistent with more recent data published by World Bank (2017). Moreover, the number reflects average compulsory schooling.
Bahrain: For 1971 and 1972, eight years of compulsory schooling are reported. However, change "will be applied in 1973/1974" (UNESCO, 1971). Following the value of the preceding years, we code a value of zero. For 1987 to 1993, zero compulsory schooling is reported. This figure contrasts with the high values before and afterward. We thus code a missing rather than a zero value. Based on the age range, 12 years of compulsory schooling are reported for 1995 to 1997. However, the compulsory program only contains six years of primary schooling with a general academic curriculum combined with religious instruction, which continues to nine years. Correspondingly, we code nine instead of 12 years for 1995 to 1997.
Barbados: There is no compulsory schooling from 1963 to 1967; however, the value for 1966 is missing. We impute this value to be zero. For the years 1995 to 1997, UNESCO yearbooks report 12 years of compulsory schooling instead of 11 in preceding and subsequent periods.
Belgium: In 1985 and 1986, UNESCO yearbooks report eight and nine years of compulsory schooling. Based on the preceding years and the age range, nine years of compulsory schooling are implausible, however. Therefore, we code eight years in 1985 and 1986.
Benin: After independence in 1960, there was a longer spell of political turbulence. In particular, several changes in power occurred at the beginning of the 1970s. According to the UNESCO yearbooks, compulsory schooling amounts to six years until 1970, zero years from 1971 to 1974, and seven years from 1975 onward. Due to the unstable nature of the government, the exact role of compulsory schooling and whether it was enforced is unclear. Therefore, we decided to code the years 1971 to 1974 as missing rather than a clean zero.
Brazil: For 1963 and 1964, UNESCO yearbooks report compulsory schooling values of four and five years. From 1965 onward, the level remains consistently at four years. Because Brazil follows the Portuguese educational system, we code a value of four for 1963 and 1964.
Brunei: For the years 1995 to 1997, UNESCO yearbooks report compulsory schooling levels of 12 years. These stand in contrast to nine years of compulsory schooling before and afterward. Because neither the educational system nor the age range of compulsory schooling changed during this period, we code nine instead of 12 years.
Cameroon: Historically, the educational system consisted of French schools in the Eastern and British schools in the Western part of Cameroon. In 1976, the British system was adopted in the entire country. We use the British system's compulsory schooling regulations throughout all periods. UNESCO yearbooks list eight years of compulsory schooling in 1969 and 1970. Given the subsequent period without any compulsory schooling, enforcement of this regulation is unlikely. We thus code a zero value for 1969 and 1970.
Canada: Compulsory schooling "[...] figures vary slightly from one Province to another" (UNESCO, 1963). Values range from seven to 10 years between 1963 and 1968 and eight to 10 years between 1969 and 1994. We take a slightly conservative view and code a value of eight years for 1963–1968 and nine years for 1969–1994.
Cape Verde: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for 1964 to 1967, compulsory schooling amounts to four years as is the case for Portugal.
Czech Republic and Slovakia: We use compulsory schooling regulations of former Czechoslovakia for both countries prior to 1994.
Egypt: For 1989 to 1991, UNESCO yearbooks report nine years of compulsory education. However, these figures are implausible given five years of primary and three years of lower secondary education. Therefore, we code eight rather than nine years. For 1995 and 1996, the yearbooks report five years of compulsory schooling. This figure does not reflect lower secondary education, which is also compulsory since the educational reforms in the early 1990s. Hence, we set the corresponding value to eight years instead of five.

Fiji: Between 1975 and 1997, UNESCO yearbooks report zero years of compulsory schooling in contrast to eight years from 1963 to 1974. We follow the World Bank convention, which codes missing values of compulsory schooling between 1998 and 2015 (World Bank, 2017). Thus, we code missing values for 1975 to 1997.
Finland: For 1967, 1971, and 1972, UNESCO yearbooks report eight instead of formerly nine years of compulsory schooling. Based on the age range and structure of the educational system, these shifts seem implausible. Thus, we code nine year of compulsory schooling.
Germany: Figures are based on West Germany prior to 1990. We code 12 rather than nine years of compulsory schooling in 1968–1970 and 1973–1988. This coding includes nine years of compulsory schooling plus an additional three years of "[...] part time vocational education" (UNESCO, 1973).
Guinea: In 1971/1972, compulsory schooling increased from eight to 12 years before it dropped again back to eight years in 1973. Throughout this period, the overall structure of the educational system remained unaltered. The only detectable change was the range of compulsory schooling from ages 7–15 to 7–22, which is implausible compared with other countries and Guinea's legal age. Therefore, compulsory schooling is coded to remain at eight years instead of 12.
Guinea-Bissau: For 1981 and 1982, UNESCO yearbooks report seven years of compulsory schooling in contrast to six years in preceding and subsequent periods. Because the educational system remained unaltered during these years, this change seems implausible. Hence, we code six rather than seven years.
Guyana: Throughout 1963 to 1997, compulsory schooling takes a value of eight years with the exception of 1981 and 1982 (nine years), 1983 (six years), and 1995 to 1997 (10 years). However, the shifts are inconsistent with the relative stability of the educational system between 1980 and 1984 and the age range of compulsory schooling from ages six to 14 for 1995 to 1997. We thus code eight years over the entire period.
India: In 1971 and 1972, the UNESCO yearbooks report various levels of compulsory schooling. In the years thereafter, only a uniform level of five years is reported. This change is justified by the fact that "[t]his information pertains to the majority of states" (UNESCO, 1975). Therefore, we also code a value of five years for 1971 and 1972.
Indonesia: In 1973 and 1974, UNESCO yearbooks report zero values for compulsory schooling. These figures stand in contrast to six years of compulsory schooling before and thereafter. Moreover, the educational system remained unaltered during this period. Hence, we code a value of six instead of zero years.
Iran: For 1966, 1967, 1973, and 1974, UNESCO yearbooks report five years of compulsory schooling in contrast to six years in preceding and intermediate periods. However, these figures seem implausible because the educational system remained unaltered during this period. Therefore, we code six rather than five years.
Iraq: UNESCO yearbooks consistently report six years of compulsory schooling. In 1983, however, five years are reported although the educational structure did not change. We code six instead of five years. Moreover, compulsory schooling is missing in 1973 and 1974. Because the educational system remained unaltered, we set the value to six years—the same as in the preceding and following years.
Israel: Between 1981 and 1987, UNESCO yearbooks report nine years of compulsory schooling in contrast to 11 years in preceding and subsequent periods. Moreover, this figure seems implausible given the age range from five to 15. Hence, we code 11 instead of nine years for this period.
Jordan: According to the UNESCO yearbooks, compulsory schooling increased from six to nine years in 1964 based on a widening of the age range. However, this increase is not observed in 1965 where the age range is again six years. Therefore, we code six instead of nine years.
Kiribati and Tuvalu: Until 1976, the islands were a British protectorate under the name Gilbert and Ellice Islands. We thus use compulsory schooling of the former protectorate for both Kiribati and Tuvalu. For the years 1975 to 1980, during which the islands became independent, we code missing instead of the reported zero values. For the years 1985 and 1986, UNESCO yearbooks report five years of schooling in contrast to nine years in the preceding and subsequent periods. Because the educational system remained unaltered during this time, we code nine instead of five years.
Kuwait: For 1982 and 1983, UNESCO yearbooks report four rather than eight years as in preceding and subsequent periods. Because the educational system remained unaltered during this period, we code eight instead of four years.
Laos: For the years 1990 to 1994, UNESCO yearbooks report eight rather than five years of compulsory schooling as in preceding and subsequent periods. Because the educational system with five years of compulsory primary schooling remained unaltered during this period, we code five instead of eight years.
Lebanon: Throughout 1963 to 1997, compulsory schooling is consistently zero years, except for 1971 where UNESCO yearbooks report a value of 12. Given the overall trend, this value seems implausible so that we code zero years.
Lesotho: The UNESCO yearbooks report compulsory schooling of eight years for the former British Crown colony Basutoland in 1964 and 1965. However, there was no compulsory schooling for the independent state of Lesotho between 1966 and 1984. Moreover, the yearbooks also report a value of zero for the colony in 1963. We thus set the value for compulsory schooling to zero for 1964 and 1965.
Malawi: For 1963 to 1965, UNESCO yearbooks report eight years of compulsory schooling based on the English schools in the former British colony. From 1966 onward, zero years of schooling are reported. Because Malawi became independent in 1964, eight years of compulsory schooling seem implausible. Hence, we code zero rather eight years.

Malaysia: From 1968 to 1984, UNESCO yearbooks report six years of compulsory schooling for some and zero or missing values for other regions. Because there is no compulsory schooling in the most populous regions, we code zero years from 1968 to 1984.
Malta: In 1986 and 1987, UNESCO yearbooks report 12 years of compulsory schooling. Based on the stable educational system, the age range, and subsequent values, these figures seem implausible. We code 10 instead of 12 years.
Mauritius: For 1981 to 1983, UNESCO yearbooks report eight years of compulsory schooling rather than seven years as in preceding and subsequent years. Because the educational system remained unaltered during this period, we code seven rather than eight years. Between 1987 and 1994, figures for compulsory schooling drop to zero. However, these values seem implausible because the educational system did not change in this period either. We code missing instead of zero values.
Monaco: For 1973 and 1974, UNESCO yearbooks report 11 years of compulsory schooling rather than 10 years as before and afterward. Because the educational system remained unaltered during this period, we code 10 rather than 11 years.
Mozambique: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for 1964 to 1967, compulsory schooling amounts to four years as is the case for Portugal.
Nauru: For 1963 to 1970, UNESCO yearbooks report nine years of compulsory schooling for European and 10 years for Nauruan schools. We code a value of 10 years.
Nepal: Historically, the Nepalese educational system consisted of English and Sanskrit schools. Until 1967, there was no compulsory schooling for either of these types of schools. Beginning in 1968, the English school system prescribed five years of schooling while attendance at Sanskrit schools was not compulsory. Following the UNESCO's convention to document compulsory schooling based on the English system from 1973 onward (UNESCO, 1973), we code five years of compulsory schooling.
New Zealand: For 1994 to 1997, UNESCO yearbooks report 11 years of compulsory schooling. However, the educational system consists of six years of primary and four years of lower secondary schooling. For this reason, we code 10 rather than 11 years. This coding choice is consistent with preceding and subsequent periods and the stability of the educational system overall.
Niger: For 1973 to 1979, the UNESCO yearbooks report compulsory schooling of 12/13 years rather than eight years as in the preceding and subsequent periods. This substantial change is not reflected in a corresponding transformation of the educational system and only represents shifts in the age range for compulsory schooling. Therefore, this extreme increase seems implausible so that we code compulsory schooling to remain at eight years throughout 1973 to 1979.
Norway: From 1968 to 1970, values of seven and nine years are reported because "[a] law passed in 1968 extended compulsory education from seven to nine years. This has been applied in most municipalities" (UNESCO, 1968).
Philippines: In 1963 and 1964, a missing value of compulsory schooling is reported. However, we decided to code a zero value because "[i]n implementation of Republic Act No. 1124, Department Order No. 1, s.1957, Article 2 states that elementary education shall ultimately be made available for all children between 7 and 13 years" (UNESCO, 1963). Hence, compulsory schooling was not yet implemented in 1963 and 1964.
Poland: Between 1963 and 1970, UNESCO yearbooks report various values of compulsory schooling. We take a conservative view and code 1963 and 1964 with a value of seven years and 1965 to 1970 with a value of eight years.
Republic of Congo: For the period 1973 and 1974, compulsory schooling dropped from an initial value of 10 to six years. From 1975 onward, compulsory schooling reverted back to a value of 10 years. Throughout this entire time, compulsory schooling age ranged from six to 16 years for boys and six to 17 years for girls. Therefore, we also code a value of 10 years for 1973 and 1974.
Romania: For 1963 and 1964, UNESCO yearbooks report seven or eight years of compulsory schooling. In subsequent years, educational regulations prescribe eight years of compulsory schooling. Based on this stability in the educational system, we set values to eight years for 1963 and 1964.
Saint Lucia: For 1985 and 1986, UNESCO yearbooks report 11 years of compulsory education rather than10 years as in preceding and subsequent periods. Because the structure of the educational system with seven years of primary and three years of lower secondary schooling did not change during these years, this shifts seems implausible. Hence, we code 10 rather than 11 years.
Sao Tome and Principe: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as is the case for Portugal.
Senegal: UNESCO yearbooks report seven years of compulsory education for 1971 and 1972 and six years for 1973 and 1974. However, compulsory primary education corresponded only to six and five years. Therefore, we code six and five years rather than seven and six.
Singapore: Compulsory schooling was only introduced in 2003. Hence, we code one missing value as zero before 2003.
South Africa: Between 1963 and 1984, UNESCO yearbooks report seven and nine years of compulsory schooling, varying by state and race. We code seven years of schooling as the corresponding figure for the black population, which constitutes approximately 80 percent of the total population.

Sri Lanka: From 1995 to 1997, UNESCO yearbooks report 11 years of compulsory schooling rather than 10 years as beforehand. Based on the age limits that remained unaltered over this period, we code 10 instead of 11 years.
St. Vincent and The Grenadines: UNESCO yearbooks report 10 years of compulsory schooling for 1968–1974 and 1978–1985 and zero years for 1963–1967, 1975–1977, and 1986–1995. Between 1996 and 2004, no values are reported. The overall structure of the educational system did not change substantially throughout all these periods so that large shifts in compulsory schooling appear implausible. We thus code values for 1963–1967, 1975–1977, and 1986–1995 to be missing rather than zero.
Suriname: UNESCO yearbooks report 11 years of compulsory schooling for the period 1995 to 1997. This figure stands in stark contrast to only six years before and afterward. Because the educational system with six years of compulsory primary schooling remained unaltered during these years, we code six instead of 11 years.
Swaziland: In the early years until 1965, the educational system consisted of European, African, and Eurafrican schools. Because education was compulsory only at European schools, which were abolished from 1966 onward, and not for the other school types, we code a value of zero.
Switzerland: According to the UNESCO yearbooks, compulsory schooling varies between seven and nine years across Swiss cantons from 1963 to 1997. In some cantons, students are additionally required to take up at least two years of "complementary part-time schooling" (UNESCO, 1963). Hence, the reported figures are likely too low. Thus, we follow the convention of UNESCO reports from 1975 to 1981 and code nine years of compulsory schooling throughout the entire period.
Thailand: In 1963 and 1964, UNESCO yearbooks report between four and seven years of compulsory schooling. Based on the age range and subsequent values, we code both observations as seven.
Tonga: For 1995 to 1997, UNESCO yearbooks report eight years of compulsory schooling. However, this figure seems implausible compared with six years in preceding and subsequent periods. Moreover, the educational system remained unaltered during these years. Hence, we code six rather than eight years. For 2012 to 2015, World Bank (2017) reports eight and then 15 years of compulsory schooling. These figures are implausible because only primary education, which requires six years of schooling, is compulsory in Tonga. Therefore, we also code six years of compulsory schooling for 2012 to 2015.
Trinidad and Tobago: In 1973 and 1974, compulsory schooling is reported to possess a value of 10 years. Before 1973 and after 1974, this figure corresponds to seven years. Because only primary schooling is compulsory with a standard duration of seven years given entry ages for primary and secondary schooling, we code a value of seven for 1973 and 1974.
Turkey: Between 1965 and 1967, eight years of compulsory schooling is reported. However, only five years of primary schooling were compulsory. In line with preceding and subsequent periods, we thus code five years of compulsory schooling.
Tunisia: From 1968 to 1981, UNESCO yearbooks report six years of compulsory schooling. For 1982 and 1983, no values are reported. From 1984 onward, compulsory schooling is documented with a value of zero until 1992. The yearbooks show 11 years of compulsory schooling for 1993/1994 and nine years from 1995 onward. The educational system consists of six years of primary schooling, three years of lower secondary schooling, and a further four years of upper secondary schooling. This structure is maintained throughout 1981 to 1995. Because zero values are implausible, we code them as missing. For 1993 and 1994, we set compulsory schooling to nine instead of 11 years.
United States: For the years 1963 to 1997, UNESCO yearbooks present values ranging from 10 to 12 years for the United States. Minimum compulsory schooling corresponds to 10 years, formally from age six to 16. Some states require students to remain in school until coming of age, implying two further years. However, there are also exemption regulations for religious groups and homeschooling. We take a conservative view and set the compulsory schooling thus to the minimum value of 10 years, which is fulfilled by all states.
Vanuatu: Historically, the educational system consists of English and French schools. Compulsory schooling years refer to regulations with respect to English schools.
Yemen: Figures are based on compulsory schooling of the former Arab Republic of Yemen and the Republic of Yemen.
Zambia: For 1963 to 1966, UNESCO yearbooks report compulsory schooling of eight years with zero years from 1967 onward. Because "[e]ducation is compulsory in certain areas only" (UNESCO, 1963–1966), we code the years 1963 to 1966 as zero.
Azerbaijan, Armenia, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan: Prior to 1992, we code compulsory schooling according to the values of the former Soviet Union. Between 1963 and 1966, UNESCO yearbooks report eight and nine years of compulsory schooling. Because primary schooling comprises only eight grades, we code eight rather than nine years.
Bosnia and Herzegovina, Croatia, Macedonia, Montenegro, Serbia, Slovenia: Prior to 1993, we code compulsory schooling according to values of former Yugoslavia. Figures of Serbia and Montenegro are taken from the Federal Republic of Yugoslavia for 1993 to 1997.