HIV/AIDS and Economic Growth: An Econometric Analysis with Particular Consideration of the Role of Education Capital Accumulation[1]

Duncan Maclaine[2]

School of Economics, University of Queensland, Australia
Australian Centre for Economic Research on Health (ACERH)[3]

Abstract: *This paper investigates, using a panel data approach, the effects of the HIV/AIDS epidemic on economic growth using a neoclassical growth model that incorporates human capital in the form of both health capital and multiple indicators of education capital. The innovations of our analysis are twofold: (i) it represents the first cross-country econometric assessment that pays particular attention to human capital accumulation through education, and (ii) it implements the dynamic panel system-GMM estimator that is know to be superior to alternative estimators that underlie previous econometric assessments of the epidemic. Based on a sample spanning 45 years and 142 countries, results indicate that the epidemic's effects on growth have been large and that a material component of this effect is due to its detrimental impact on the accumulation of education capital. For the full sample and a sub-sample of developing world countries, the impact of a 1% increase in adult HIV prevalence is estimated to be a reduction in income per capita of between 0.12% and 0.16% on average, which is substantially larger than that found in previous econometric assessments of the epidemic.*

JEL Classification: O11; O15; O47; I12; I20.
Keywords: HIV; AIDS; Growth; Human Capital; Health; Education

---

[1] Presented at the New Zealand Association of Economists (NZAE) conference in Christchurch on 27 June 2007. This paper is based on work leading to a PhD thesis and is an entry for the NZAE Jan Whitwell prize.

[2] School of Economics, University of Queensland, Brisbane, QLD 4072, Australia; Telephone +61 7 33654028. Email: d.maclaine@uq.edu.au

[3] Financial support from a University of Queensland Joint Research Scholarship and an ACERH research scholarship are gratefully acknowledged.

1. Introduction

It is now twenty-seven years since the first clinical evidence of AIDS emerged. The death and suffering caused by the epidemic during these years represents a monumental human tragedy. This paper seeks to determine if we are also witnessing the unfolding of an economic tragedy.

It is estimated that between 34 and 47 million people worldwide were living with HIV in 2006 (UNAIDS, 2006). Recent data indicates that HIV/AIDS is the fifth leading cause of mortality worldwide, fourth in low and middle-income countries and the number one cause of death in Sub-Saharan Africa, closely followed by malaria (Lopez *et al.*, 2006).

Prevalence rates of HIV worldwide appear, on average, to be stabilizing, with some of the most affected regions in Sub-Saharan Africa experiencing some declines, although this trend is neither widespread nor particularly strong (UNAIDS, 2006). Table 1, included in Appendix A, details prevalence estimates for the years 2005 and 2003.[4] Swaziland continues to experience the worst epidemic, with one-third of 15 to 49 year-olds infected with HIV. Outside Africa, Haiti has the highest prevalence rate of 3.8%. Although not featured in Table 1, India has the largest absolute number of adult HIV cases in the world, approximately 5.6 million, despite a prevalence rate of less than 1%, due to the sheer size of its population.

Readers who are interested in a detailed assessment of the epidemic are referred to the comprehensive report published by UNAIDS in conjunction with the WHO entitled 'Report on the Global AIDS Epidemic 2006.[5]

HIV/AIDS has now been in existence long enough to attempt a systematic assessment of the economic growth impact of the epidemic using panel data. A small number of recent econometric studies have investigated the role HIV/AIDS has on determining cross-country differences in income per capita and / or its role in the evolution of a country's economic performance over time. Each study makes use of a variant of the Solow growth model that explicitly incorporates human capital as a factor of production. These papers include Bloom and Mahal (1997), Dixon *et al.* (2001), Papageorgiou and Stoytcheva (2004), Tandon (2005), and McDonald and Roberts (2006).

The analysis presented in this paper represents the first cross-country econometric analysis that pays particular attention to the role HIV/AIDS plays in human capital accumulation

[4] Prevalence estimates for 2005 and all previous years have been substantially revised downwards by UNAIDS/WHO as a result of improved estimation methods. For example, in Swaziland, HIV prevalence was previously estimated to be 38.8% in 2003, whilst latest estimates for 2003 are 32.4%. For Cambodia, previous 2003 prevalence estimates were 2.6%, whilst the updated estimate for 2003 is 2.0%.

[5] This report is available at www.unaids.org

through education. This innovation is important because the role the epidemic plays in the accumulation of human capital through education has, to date, largely been investigated through calibration and simulation of theoretical overlapping generations (OLG) or computable general equilibrium (CGE) models, with such techniques open to the criticism that they rely heavily on assumptions that can often be difficult to justify.

Another innovation of this paper is the use of the dynamic panel system-GMM estimator of Arellano and Bover (1995) and Blundell and Bond (1998). This estimator has been shown to be superior to the more traditional difference-GMM estimator of Arellano and Bond (1991) under the precise circumstances at hand here. It is argued that these developments in the econometric literature cast doubt over the accuracy of earlier estimates of the macroeconomic impact of HIV/AIDS obtained from growth regressions that have employed the difference-GMM estimator**.**

This role HIV/AIDS plays in the accumulation of education capital has been the focus of much of the very recent economic literature. Ferreira and Pessoa (2003), Corrigan *et al.* (2005) and Bell *et al.* (2006) all highlight, through the use of calibrated overlapping-generations models, the potentially large negative impact HIV/AIDS may have on human capital accumulation through education and in turn on future income per capita. Arndt (2006) also highlights significant negative impacts based on a CGE model for Mozambique. Young (2005) also considers the role of education capital accumulation, but in addition, assesses the potential role HIV/AIDS has on the determination of fertility rates.

One might expect that as the body of literature assessing the economic growth implications of HIV/AIDS grows we might be converging towards a consensus view, but in fact, the opposite appears to be occurring. Bell *et al.* (2006), focussing on South Africa, conclude that if the epidemic continues unabated, in the absence of government policy to tackle the epidemic, there could be a virtual economic collapse. In contrast Young (2005), also focussing on South Africa, concludes that by also considering the role the epidemic has on fertility rates, future per capita income may, in fact, rise.

In the light of such contrasting assessments, it is considered a worthy exercise to conduct a rigorous empirically investigation of the impact HIV/AIDS has on welfare by carefully considering the role of human capital accumulation. Specifically, in this paper the role human capital accumulation, through both health and education, has on the welfare, as measured by GDP per capita, is examined.

The analysis is fortunate to benefit from a number of recent improvements in coverage and quality of data, allowing for an increased cross-section sample size and potentially less measurement error associated with the included variables. For example, the latest time series estimates of HIV prevalence from UNAIDS (2006), represents an improvement on previously available data due to better estimation methods and improved data availability. Additionally, the incorporation in this study of an additional five years of data for which HIV/AIDS has been in existence, relative to the other recent cross-country empirical work, provides the potential for significant improvements in the reliability of estimates of the effects of HIV/AIDS prevalence on economic growth.

This paper is organised as follows: In section two, a summary of various aspects of the relevant literature is presented. In section three, as a precursor to a discussion of the data-set, the problems surrounding the measurement and specification of human capital is presented. Section four discusses the available data, presents the economic and econometric models and introduces the methodology employed. Section five reports and discusses the results and section six submits these results to rigorous robustness tests. Finally, section seven summarises and concludes.

2 Literature

A summary of various literatures, relevant to the task at hand, is presented here. Firstly, an overview of the concept of human capital and its importance for economic growth is presented. Secondly, the way in which human capital has traditionally been incorporated into empirical growth models is discussed. Thirdly, a review of some of the more influential macroeconomic research into the HIV/AIDS epidemic is presented. Lastly, a separate discussion of previous cross-country econometric studies is provided.

2.1 The concept of human capital and its importance for growth

"In economic terms, health and education are the two cornerstones of human capital, which Nobel Laureates Theodore Shultz and Gary Becker have demonstrated to be the basis of an individual's economic productivity" (WHO, 2001 p.21).

Modern human capital theory, to which Shultz and Becker were the main contributors, sees education and skills as capital goods for which acquisition is costly and where individuals invest in it with a view to increasing their productivity and therefore incomes.

Although many theorists acknowledged that health, in addition to education, is an important component of human capital, Grossman (1972) was the first to construct a theoretical model of the demand for health capital. Grossman described the demand for health and health care via the theory of human capital. Health cannot be purchased from the market like many other goods and services but is instead 'produced' by the individual through the consumption of health inputs (such as services). Health care services are an input into the model with individuals not demanding health care services in particular, but demanding the output of these services - better health. In the Grossman model, people inherit an initial stock of health capital which depreciates over time and can be increased with health investments. Healthy workers are more physically and mentally able and therefore more productive. A healthier population results in less days lost through illness and individuals remain in the workforce longer, increasing incomes. This notion is described in Grossman's model as the 'production benefits' of health.[6] Causality is not just in one direction, however – higher incomes increase the resources available for individuals to invest in their own health. Additionally, a healthier population, by increasing life expectancy, may also generate greater savings.

Education, according to human capital theory, is regarded as a prerequisite for an individual to obtain productive skills, enabling individuals to improve labour productivity and income. In the context of a neoclassical production function, this could be considered a 'direct effect', where education is treated as an additional input in the production process. A number of researchers also stress indirect effects of education on growth. Education can lead to improvements in technology as a result of innovation and research, improved processes and products or an increased ability for countries to adopt technological innovations from elsewhere (Romer, 1990; Aghion and Howitt, 1998; Nelson and Phelps, 1966; Hall and Jones, 1999). Macroeconomic models emphasizing this indirect role are typically referred to as endogenous growth models.

As in the case of health, causality runs in both directions, with higher incomes providing greater resources for the individual to invest in their own or their children's education.

Health and education should not be viewed as mutually exclusive components of human capital – they are interlinked in many ways. Grossman (1972) suggests that individuals with higher levels of education are more efficient in addressing their health needs, whilst Schultz (1999) stresses that better health enables an individual to utilize better the knowledge and skills they acquired through education. Improved health, by increasing life expectancy, can

[6] In addition to this 'production benefit', individuals gain satisfaction from simply being healthy – a notion defined by Grossman as a 'consumption benefit' of health.

increase the incentive to invest in one's own education due to the greater possible returns from the investment (Barro, 1996; Ferreira and Pessoa, 2003). Corrigan *et al.* (2005) highlight generational linkages between health and education by arguing that the children of adults, who drop out of the workforce due to ill-health, may enter the labour-force prematurely, thereby forgoing their education to maintain the household's income.

2.2 Human Capital in Empirical Growth Models

In the traditional neoclassical growth model, the classic Solow (1956) model, output is modelled as a function of physical capital and labour only, and human capital variables are not considered. This model did not do a very good job at explaining income disparities between countries. In their seminal 1992 paper, Mankiw Romer and Weil (1992), henceforth MRW, proposed adding human capital, in the form of education, as an additional factor of production. MRW (1992) concluded that this modification removes the bias in the coefficient estimates of physical capital and population growth that would otherwise be present and that human capital helps to explain cross-country income differences. It results in a reduction in the estimated effect of physical capital accumulation and population growth to levels that reconciled more closely with much of the empirical evidence. This 'Augmented Solow model' allows for the 'direct effect' of education on growth highlighted above. Endogenous growth models attempt to capture the possibility of 'indirect effects' of education on growth.

Islam (1995) extended the MRW (1992) framework by introducing a panel data approach. The approach did not reveal a significant relationship between human capital, in the form of education, and growth. This was a worrying result given the wealth of microeconomic evidence of the significant returns to human capital as well as strong evidence from the early growth regression literature. The traditional focus of earlier empirical work on cross-country macroeconomic growth was on the importance of education in improving human capital. Knowles and Owen (1995) extended the MRW (1992) model by including health capital as well as education capital. Their results confirm the importance of health for economic growth, however the relationship between education capital and growth is found to be insignificant. This result is common in studies that incorporate both health and education capital, including those investigating the role of HIV/AIDS on economic growth discussed in Section 2.3. Although health is commonly accepted as being an important component of human capital, the majority of empirical growth research continues to focus purely on education as a measure of human capital.

In a comprehensive critical summary of the empirical literature on the impact of human capital (in the form of education) on the economy, Sianesi and van Reenen (2003), highlight a number of factors that may have contributed to the often disappointing results highlighted above. These include; the selection of poor or inappropriately narrow choice of proxy for human capital, data measurement error and changes in classification criteria, and other methodological problems such as not controlling for potential reverse causality. In this paper, an attempt is made to address some of these common criticisms.

2.3 HIV/AIDS and the Economy

Macroeconomic analysis the HIV/AIDS epidemic typically focuses on its impact on GDP or GDP per capita. Many consider such a measure to be a less than perfect indicator of welfare, however absent a consensus on an alternative, it continues to be the main focus. A review of the literature involving the macroeconomic analysis of HIV/AIDS highlights that any estimated effect of HIV/AIDS differs according to (i) the economic framework and modelling approach utilised, (ii) the country or countries of focus, (iii) the impact channels that are accommodated and the assumptions underlying the magnitude of these, (iv) the time frame of the analysis, and (v) the epidemiological estimates underlying the economic projections.

Earlier macroeconomic studies relied on HIV prevalence projections that turned out to be serious under-estimates of the scale of the epidemic. It may also be argued that these early studies were somewhat narrow in terms of the impact channels that the economic framework allowed.

Impact channels of HIV/AIDS

A summary of the individual channels in which HIV/AIDS can affect economic outcomes is presented here along with a brief discussion of some of the more influential research in the macroeconomic literature.[7] The purpose of this section is not to report in extensive detail specific results from studies, but rather to highlight in general the impact channels considered by them and their broad conclusions. It is, however, worthwhile to present a more detailed discussion of the recent wave of studies that consider the role HIV/AIDS plays in education capital accumulation.

[7] Those studies involving cross-country econometric analysis are discussed separately in Section 2.4

A - Productivity of HIV/AIDS workers: Workers with HIV/AIDS, and especially those without access to antiretroviral therapies, are likely to be less productive in the workplace and take more days off work due to illness. Depending on the structure of the labour market, this may result in lower wages for those individuals.

B - Savings behaviour of households directly affected by HIV/AIDS: The saving rates of households affected by HIV/AIDS may fall due to the need to spend more of their income on health care. Additionally, savings will decline if household wages have fallen due to the individual being less productive at work or dropping out of the workforce completely.

C - Average skill-set of the workforce: As HIV/AIDS victims drop out of the workforce, their replacements, assuming the existence of available labour, are likely to bring less experience and skills into the role.

D - Population and labour supply: Deaths due to HIV/AIDS will lead to a lower population than in the absence of the epidemic. The size of the labour force will also decline due to morbidity and mortality of working age individuals. Additionally, individuals may drop out of the labour force to care for sick spouses or relatives. There may also be absenteeism due to employee attendance of funerals (Liu *et al.*, 2004). The proportional reduction in the size of the labour force as a result of HIV/AIDS may be larger than the reduction in overall population due to the epidemic disproportionately affected young adults. In a neoclassical growth framework this impact channel, taken in isolation, has the effect of increasing GDP per capita due to a higher capital-to-labour ratio.

The above four linkages between HIV/AIDS and macroeconomic outcomes were the main impact channels addressed in the first wave of studies that emerged in the early 1990s. Two of the influential early studies by Cuddington (1993a) and Cuddington and Hancock (1994) utilized a neoclassical, single-sector Solow model to analyse the effect of HIV/AIDS on growth in Tanzania and Malawi. Assumptions were made about the share of medical expenses paid out of household savings, demographic projections, how less productive an HIV/AIDS worker is relative to a healthy worker, and the reduction in average experience embodied in the workforce. The models were calibrated and simulated over a range of plausible assumptions. The general conclusion was that there would be a modest negative impact on GDP growth per capita.[8] These early papers were criticised on the basis that they didn't take

[8] Where it was assumed that HIV/AIDS would have no impact on household savings rates, the models produced a small positive effect on GDP per capita.

into account the fact that these African countries had significant levels of under-or-unemployment. The existence of surplus labour could raise the possibility that models assuming full-employment could overstate the economic impact as HIV/AIDS victims can be replaced in the workforce with healthy unemployed or under-employed. In response to these criticisms, Cuddington (1993b), Cuddington and Hancock (1995) moved from a single-sector Solow model to dual-sector (formal and informal labour market) Solow model with surplus labour. The results derived from the dual-sector models, however, are similar to that of the single-sector models. Over (1992) also considers a dual-sector model, with the labour market disaggregated by productivity and between urban and rural sectors. Over (1992) makes the assumption that HIV/AIDS is most prevalent amongst the most productive sectors of the economy. Applying this framework to thirty African countries over the period 1990-2025, a modest negative impact on GDP per capita is found. Other early studies include the application of a detailed CGE model for Cameroon by Kambou *et al.* (1992). HIV/AIDS was incorporated by introducing an exogenous shock to the size of the labour force and an increase in public and private health expenditures resulting in an estimated halving of the GDP growth rate between 1987 and 1991.

In retrospect, these early studies suffer from deficiencies in relation to the severe underestimates of the scale of the epidemic that were employed and that they were arguably narrow in the range of economic impact channels that were incorporated.

E - Public sector: Public sector health care costs will most likely rise, placing pressure on the government budget. A comprehensive analysis of other pressures on the public sector as highlighted by Haacker (2002a) include; increased social security costs, increased death benefits and sick leave in the public service, reduced contributions to public sector pension schemes that involve guaranteed levels of benefits, and the need to increase the number of school teachers recruited and trained due to deaths in the profession. Because HIV/AIDS disproportionately affects the working age population, the government's tax base will decline and increase the dependency ratio as mortality and morbidity in the workforce increases.

***F - Business operating environment, total factor productivity and technological progress*:** A number of microeconomic studies have investigated whether the HIV/AIDS epidemic has added significantly to business costs. In a review of published studies in this area, Liu *et al.* (2004) conclude that although the literature is far from satisfactory, there is a consensus view that HIV/AIDS has increased costs due to, among others things, increased sick leave benefits,

higher life and health insurance premiums (if borne by the employer), pensions, funeral benefits, lower productivity, and the higher costs associated with recruiting and training new staff to replace sick employees.

Arndt and Lewis (2000) utilise a disaggregated dynamic CGE model to estimate the impact of HIV/AIDS on economic growth prospects in South Africa between 2001 and 2010. In addition to impact channels A through D, the analysis incorporates the role the epidemic plays in the determination of total factor productivity (TFP) in the form of increased hiring and training costs, staff absenteeism, and a slowdown in technological adoption. The consideration of TFP was absent from the earlier CGE analysis by Kambou *et al.* (1992), as well as the earlier studies utilising calibrated neoclassical growth frameworks. Arndt and Lewis (2000) also undertake a more detailed treatment of the impact HIV/AIDS has on the government sector. Incorporation of a time dimension in the modelling, absent from Kambou *et al.* (1992), allows for potential cumulative effects of lower investment and TFP growth to be captured. A significantly greater negative impact of HIV/AIDS on GDP and GDP per capita is found relative to earlier studies due to a combination of more up-to-date demographic projections and also the significant role TFP and the government sector plays in driving the results. Over the simulation period, GDP growth declines by as much as 2.6% per annum and GDP per capita by 2010 is up to 8% lower than it would be in the absence of HIV/AIDS.

G - International competitiveness and investment: In the event that HIV/AIDS leads to reductions in worker productivity, a slowing in the rate technological innovation, or rises in business operating costs, a country that experiences a more severe epidemic than its competitors could lose any comparative advantage it held, resulting in a reduction in exports and foreign direct investment (FDI). For an open economy, any increase in the capital to-labour ratio as a result of the epidemic could lead to a decline in the return to capital that discourages domestic investment, increases capital outflows and reduces FDI (Haacker, 2002a).

Two studies focusing on Botswana by MacFarlan and Sgherri (2001) and BIDPA (2000) use calibrated dual-sector variants of neoclassical growth models to estimate medium to long-term growth prospects. Both papers consider impact channels A through E as well as F and G in the form of TFP growth and the role of international capital flows. MacFarlan and Sgherri

(2001) undertake a range of simulations with varying assumptions that produce negative impacts which are broadly similar in magnitude to the South African study by Arndt and Lewis (2000), with the largest impacts occurring due to reduced labour productivity and the cumulative effects of slower capital accumulation due to greater HIV/AIDS-related health care expenditure. The analysis undertaken by BIDPA (2000) produces significantly smaller estimates of the epidemic's impact on GDP growth and virtually no effect on GDP per capita. This difference is mainly driven by the assumption in BIDPA (2000) that any reduction in domestic savings will be offset by increased capital inflows into the country. This is quite a strong assumption as it is quite likely that although some of the domestic shortfall in savings may be offset by inflows, capital flight should be considered a distinct possibility especially if the epidemic severely impacts the availability and reliability of the skilled labour pool.

In contrast to the majority of studies that predict negative effects of HIV/AIDS on per capita GDP, albeit small effects in some cases, a report focusing on South Africa produced by the Bureau of Economic Research in Stellenbosch (BER, 2001) predicts that HIV/AIDS prevalence will result in future increases in per capita GDP. Focusing on the period from 2000 to 2015, a comprehensive macroeconometric model is utilized that incorporates the effects of impact channels A through F. The positive impact on GDP per capita is a consequence of the strong demand-side focus of the model with large HIV/AIDS related expenditures, both public and private, contributing to aggregate demand and hence GDP. Additionally, it is assumed that capacity utilization increases as labour losses due to HIV/AIDS can be easily replaced out of a large pool of unemployed individuals, especially in the unskilled sector. The BER's modelling does point to a significant reduction in potential GDP however, that could lead to a reversal of the positive impact on GDP per capita beyond the horizon considered in the analysis.

H - Social capital: HIV/AIDS may affect macroeconomic outcomes through the erosion of social capital. Bonnel (2000) argues that HIV/AIDS reduces social capital by: worsening fiscal deficits and reducing the ability of government to provide social services, eroding social networks and support mechanisms in local communities, reducing the macroeconomic management capacity of government, and reducing the ability of government to maintain an effective regulatory environment and legal framework due to a reduction in the skilled labour force. These factors are hypothesized to reduce GDP per capita because they result in "increased transaction costs and reduced efficiency in production" (Bonnel, 2000, p.829).

I - Incentives to invest in skills and training: An employer facing high levels of staff turnover due to the prevalence of HIV/AIDS may have reduced incentives to invest in additional training and skills for its staff. Individuals with HIV/AIDS would also face reduced incentives to invest in advancing their own skills.

Haacker (2002a) undertakes a detailed analysis of the impact of HIV/AIDS on the economies of Southern Africa. The role of HIV/AIDS on the public sector, formal education sector, the workplace, and the training, experience and productivity of the labour force are explored in detail. In the formal modelling of the impact on per capita GDP, a growth accounting exercise is undertaken using a version of the simple neoclassical growth framework. Although Haacker (2002a) discusses the potential impact of HIV/AIDS on formal schooling and the incentives for employers to invest in on-the-job skills training, the modelling is confined to the effect HIV/AIDS has on the average experience of the labour force due to the impact the epidemic has on the age profile of workers. Other impact channels that are incorporated into the modelling are the effect HIV/AIDS has on business costs and TFP, population and labour supply, and savings and capital accumulation.[9] A closed-economy version of the model is calibrated and simulated with key parameters informed by available microeconomic and sector specific data. Despite declines in TFP, labour force productivity (due to lower average experience of the workforce) and private savings, GDP per capita is expected to increase in the long-term due to the effects of lower population growth and an increased capital-to-labour ratio. Estimates range from an increase of 3.9% in Mozambique to 9.6% in Botswana. Once the analysis is extended to an open-economy framework, the increase in the capital-to-labour ratio results in lower returns to capital, declines in FDI, and increased capital outflows. These added effects reverse the findings from the closed-economy model with an estimated decline in GDP per capita of 2% predicted, on average, in the long-term across Southern Africa.

J - Life expectancy and savings: Separate from the effect HIV/AIDS may have on savings rates due to increased health care expenditure, the near certain premature death of individuals with HIV/AIDS, especially when expensive antiretroviral therapies are unavailable, may itself lead to lower rates of savings. This is because, according to the life cycle or permanent-

[9] Although Haacker (2002a) presents a detailed discussion of the impact HIV/AIDS may have on the public sector, this is not explicitly incorporated into the formal modelling.

income hypothesis, the reason individuals save is for future consumption. This possibility is allowed for by Corrigan *et al.* (2005). It is also possible that even those individuals who do not currently have HIV/AIDS might save less in the face of widespread prevalence of HIV/AIDS due to the lower average life expectancy of the population.

K - Accumulation of education capital: Widespread prevalence of HIV/AIDS, by reducing life expectancy, can reduce the incentive to invest in one's own education due to the lower possible returns from the investment (Ferreira and Pessoa, 2003; Bell *et al.*, 2006). Additionally, the incomes of those households where an adult is directly impacted by HIV/AIDS are likely to decline, therefore reducing the ability to invest in their children's education. This could lead to children entering the labour market prematurely to supplement household income. The result is a vicious cycle where these children, because of their lower levels of education find themselves in lower-skilled, low-paid jobs, and therefore have limited resources to invest in their own children's education (Corrigan *et al.*, 2005; Bell *et al.*, 2006). Children who become orphans are also likely to struggle to continue in formal education relative to non-orphans, at least beyond the level in which the state will provide resources. According to Haacker (2002a), the number of orphans as a result HIV/AIDS in, for example, Botswana, could be as high as 7.2 per cent of the population by 2010.

L – Fertility: Young (2005) suggests two ways in which HIV/AIDS could affect the decision to have children. Firstly, fertility rates might decline due to the unwillingness of individuals to engage in sexual activity because of the risks of contracting HIV/AIDS. Secondly, deaths of workers due to HIV/AIDS can lead to labour scarcity and therefore higher wages. This increase in wages could tempt more women into the labour force and reduce women's demand for children due to the higher opportunity cost of time.

The economic impact of HIV/AIDS through its ability to slow the accumulation of human capital adversely has been the focus of much of the recent wave macroeconomic literature. The potential for the epidemic to slow economic growth through its effect on education was not considered in the literature until very recently. One such study, by Ferreira and Pessoa (2003) develops a theoretical OLG model whereby the long-run economic costs of HIV/AIDS are driven by a reduction in the incentive to invest in one's own education due to the epidemic's impact on life expectancy. Simulations with their model that focus purely on this 'life expectancy' channel produce significantly larger estimates than earlier studies of the macroeconomic impact of HIV/AIDS. GDP per capita in the long-run steady state is

estimated to be 26.3% lower than in a no-AIDS counterfactual scenario for the nine countries with the worst epidemics.

Corrigan *et al.* (2005), also utilize an OLG framework to consider the effect of the creation of orphans on the accumulation of education capital and its transmission across generations. The incomes of those households impacted by HIV/AIDS are negatively affected by morbidity and eventually mortality, reducing the ability to invest in a child's education. This can lead children into entering the labour market prematurely. A vicious cycle ensues as those children obtain lower levels of education and hence lower-skilled, low-paid jobs, and therefore have limited resources to invest in their own children's education. The life expectancy effect on human capital accumulation proposed by Ferreira and Pessoa (2003) is not considered by Corrigan *et al.* (2005), however its effect on saving and physical capital accumulation is captured. Calibrating the model for a typical sub-Saharan African country and assuming prevalence rates of 20% for one generation, their baseline scenario suggests that HIV/AIDS will lower income per capita by 6.25% in the long-run.[10] Similarly, Bell *et al.* (2006) focus on these intergenerational effects utilising an overlapping-generations framework, but in addition consider the 'life expectancy' effect proposed by Ferreira and Pessoa (2003). The model generates much larger negative long-run predictions than any previous studies of its kind. Applying the model to South Africa, if the epidemic continues unabated, in the absence of government policy to tackle the epidemic, economic collapse arises within two generations due to the cumulative effects of a breakdown in transmission of human capital across generations. By contrast, their modelling also suggests that government policies aimed at combating the disease and its economic effects can prevent such a collapse, albeit with a huge fiscal burden in the order of four per cent of GDP. Arndt (2006) investigates the growth impact HIV/AIDS may have on the Mozambiquen economy by paying special attention to the demand and supply-side impacts on the education sector and human capital accumulation. Simulations are conducted using a disaggregated CGE model incorporating estimated education and human capital transition matrices. Simulations over the 1997 to 2010 period suggest that HIV/AIDS could reduce real GDP by up to 20 per cent, relative to a no-AIDS baseline, with the impact on human capital, and in turn human capital's role in determining technological progress being a major contributor. By 2010, the GDP growth rate is estimated to be 4.3% lower in the baseline scenario relative to the no-AIDS scenario. This negative impact is shown to be materially reduced by appropriate policy initiatives aimed at the education sector. Impacts on per capita measures are not as large due to population effects:

[10] The negative impact becomes 14 per cent when HIV/AIDS is assumed not to be eradicated for two generations.

average per capita growth rates are one per cent lower average per capita growth rates and per capita GDP is 12 per cent lower by 2010.

Not all recent research points to large negative impacts of the HIV/AIDS epidemic on economic growth. In the provocatively-titled paper "The gift of the dying: The tragedy of AIDS and the welfare of future African Generations", Young (2005) concludes that the epidemic will raise future per capita GDP. A novel theoretical framework is utilized that embeds a Beckerian model of household behaviour into a Solow-style macroeconomic framework. The model endogenises labour force participation, fertility and education decisions. Calibrating and simulating the model for the South African economy, it is found that although there is a large detrimental impact on the accumulation of human capital, this is more than offset by a 'fertility effect'. HIV/AIDS lowers fertility rates due to the unwillingness of individuals to engage in sexual activity, and also due to a reduction in women's demand for children as they are tempted into the labour force due to relatively high real wages. Mortality resulting from the AIDS epidemic and the unwillingness to engage in sexual activity reduces the supply of labour over time and consequently increases wages. Lower population growth, in the context of a Solow model with exogenous saving, increases future per capita income.

2.4 The Econometric Analysis of HIV/AIDS and Economic Growth

A small number of cross-country econometric studies have attempted to identify what effect HIV/AIDS has on economic growth. The first such study by Bloom and Mahal (1997) is based on cross-sectional regressions estimated using a variant of the well known MRW (1992) growth model. Estimation was based on a sample of 51 countries for the period 1980 to 1992. No statistically significant effect of adult HIV/AIDS prevalence or incidence on economic growth was detected. Many have subsequently questioned these results on a number of grounds; they were based on serious underestimates of HIV/AIDS prevalence, it may have been too early for the impact of the epidemic on morbidity and mortality and in turn economic activity to be detected, and that many countries experiencing severe epidemics were excluded from the sample (Dixon *et al.*, 2001; McDonald and Roberts, 2006). Additionally, the methodology used has been questioned. In the subsequent analysis of Dixon *et al.* (2001), Tandon (2005) and McDonald and Roberts (2006), henceforth MR (2006), a theoretical derived relationship between HIV prevalence and health capital is determined through a

separate model, rather than HIV/AIDS variables being added in an *ad hoc* way directly into a growth equation as is the case in Bloom and Mahal (1997).

Papageorgiou and Stoytcheva (2004) utilize an augmented Solow model where AIDS incidence in the population is included as an extra variable to the well known MRW (1992) specification. The analysis is based on a sample of 89 developed and developing countries over the period 1979 to 2000. Based on both cross-sectional and panel methods and a variety of alternative specifications, an increase in AIDS incidence is found to have a significant and negative impact on GDP per worker for the full sample and a sub-sample of non-OECD countries. Dixon *et al.* (2001) make use of an augmented Solow specification estimated as a two-way 'fixed effects' model. The dependent variable is growth in income per capita and included in the regressors are both health and education capital.[11] A separate equation is specified for health capital, proxied by a measure of life expectancy. The income equation and the health capital equation are estimated as a system using the seemingly unrelated regression (SUR) method.[12] Based on a sample of 49 African countries, no significant impact of HIV prevalence on economic growth is detected.

However, splitting the sample into 'Southern & Eastern Africa' and 'Rest of Africa', results in a negative and statistically significant impact of HIV prevalence on life expectancy and in turn on economic growth in the 'Rest of Africa' sub-sample. Counter-intuitive results are found for the Southern & Eastern Africa sub-sample for which the authors attribute to a breakdown of 'normal economic relationships' in those countries with extremely high prevalence rates. There are some methodological concerns involving the analysis of Dixon *et al.* (2001). Firstly, the health equation estimated is static, that is, past levels of health capital are not included as regressors. Conceptually, it would be preferable to employ a dynamic specification since the model concerns the stock of, not investment in, health. Secondly, the counter-intuitive results for the Southern and Eastern Africa sub-sample may be more to do with the small sample size (n=16), rather than the breakdown of 'normal economic relationships'.

[11] This represents a major point of distinction from the analysis of Bloom and Mahal (1997) and Papageorgiou and Stoytcheva (2004) where only education capital was included in the model along with AIDS incidence or HIV prevalence being added directly to the growth equation as an extra regressor.

[12] The SUR method was defended by the authors on the grounds that in the income equation, the dependent variable is the *growth rate* in income per capita, whilst the *level* of income per capita is included as one of the regressors in the health capital equation. As such, the system is not simultaneously determined. In this study, and in the studies of MR (2006) and Tandon (2005), the income equation is specified in terms of levels, rather than growth rates.

The two most recent studies that most closely resemble the analysis in this paper are those of MR (2006), and Tandon (2005). Both papers conclude that a 1% rise in HIV prevalence leads to between 0.05% and 0.08% decline in income per capita based on full world or developing world samples.

Both papers utilize a version of the augmented Solow model with the dependent variable in the growth equation being income per capita and health and education capital included as regressors. A separate equation is specified for health capital. In both cases the models are estimated using five-yearly averaged panel data. In MR (2006), the health capital equation is estimated using dynamic panel difference-GMM estimation of Arellano and Bond (1991), a form of instrumental variable estimation. In both studies, the predicted values for health capital are then used as instruments for health capital in the growth equation, which in turn is specified with two-way fixed effects and estimated using the difference-GMM estimator.[13] Statistically significant and material negative effects of increased HIV prevalence are detected in both studies for the full sample and in MR (2006) for a sub-sample of developing and African countries. There are some similar concerns regarding methodology as for the study by Dixon *et al.* (2001). Firstly, some disappointing results are obtained for the Latin American & Caribbean sub-sample in MR (2006) and for the Asian sub-sample in Tandon (2005) – most likely due to small sample size. This problem could potentially be overcome by the use of a full set of intercept and interaction dummy variables as a means of sub-sample analysis, improving the efficiency of the estimation. Secondly, Tandon (2005) employs a static specification of the health capital equation and, for reasons outlined above, a dynamic specification may be considered more appropriate. MR (2006) employ a dynamic specification of the health capital equation.

A general concern with these econometric studies is that, in each of them, a proxy variable for education capital, such as secondary enrolment rates or average years of schooling in the adult population, is included in the growth equation and found to be statistically insignificant or the opposite sign to what would be expected. As discussed in Section 2.2, this result is common in the broad cross-country growth literature.

[13] As discussed in Section 3, this paper adopts the system-GMM estimator of Blundell and Bond (1998) which is considered *a priori* superior to the difference-GMM estimator of Arellano and Bond (1991).

3 Specifying Human Capital

3.1 The measurement of human capital

Since human capital is a latent variable, in the sense that it is (like utility) immeasurable, researchers are compelled to find 'indicators' of or proxy variables for, the stock of the latent variable, the selection of which is a difficult task. There is often a trade-off between a measure that is conceptually appealing but has a limited time-series or country availability and a measure that is less appealing but is widely available over time and across countries. In empirical studies many different proxies for education as a measure of human capital have been utilized. Wossmann (2003) provides an excellent analysis of the many proxies used in the growth literature. Historically, the most common proxies have included school enrolment ratios (even though this is conceptually a flow measure), average years of schooling and indicators of cognitive ability (e.g. adult literacy rates and other internationally comparable tests of cognitive achievement). In recent years, measures of average years of schooling in the adult population have become the most popular indicators of the stock of human capital for empirical work (Wossmann, 2003). The use of adult literacy rates has largely fallen out of favour as it become clear that it represents a very noisy measure of human capital: adult literacy rates disregard the level and type of literacy and the acquisition of skills beyond basic literacy (Wossmann, 2003). The school enrolment ratio, as a proxy or indicator for the stock of human capital, has also come under criticism recently. Enrolment ratios represent a flow variable and pertain mostly to people who are not currently in the labour force.[14]

Although measures of educational attainment are regarded by many as an improvement on the use of literacy rates and enrolment ratios as a proxy or indicator for human capital, this measure still has some major shortcomings: it does not allow for the diminishing returns to education or for the quality of education received (Wossmann, 2003). Wossman (2003) proposes a quality-adjusted measure of human capital that captures cross-country differences in the quality of education and also diminishing returns to education. Although it would be preferable to utilize such an indicator in the present study, its time and cross-sectional coverage is insufficient for our purposes.

In this study, both educational attainment in the adult population and enrolment ratios were initially considered for inclusion in the model. Although educational attainment is

[14] Hanushek and Kimko (2000) and Wossmann (2003) are also critical of its use a proxy or indicator for the flow of human capital investment.

conceptually closer to the notion of the 'stock' of education capital, enrolment ratios were eventually chosen for inclusion as it had a wider country coverage. Notwithstanding the conceptual problems regarding the use of enrolment ratios as an indicator or proxy for the stock of human capital, some comfort may be drawn from the observation by Wossmann (2003, p. 258) that "enrolment ratios may not be an altogether bad proxy for the quantity of schooling after all". This observation was made after discovering the high degree of correlation between the series measuring educational attainment in the adult population and school enrolment ratios.

In the studies concerning the impact of HIV/AIDS on cross-country economic growth, both enrolment ratios and educational attainment have been used.

Turning to health as a component of human capital, a number of alternative indicators or proxies for the population's health status are available. The most commonly-used include; life expectancy at birth (or its shortfall from a benchmark age), infant mortality rates, and under-five mortality rates. Life expectancy and mortality rate data are widely available and both are considered in this study. Both indicators do, however, have some shortcomings. Life expectancy statistics do not capture the number of sick individuals or the importance of the morbidity effects of disease, including HIV/AIDS (MR, 2006). Similarly, infant or under-five mortality rates only measure the impact of fatal diseases. MR (2006) argue, however, that infant mortality rates will partly reflect the general quality of health of the population, as infants are the most likely to suffer from deterioration in the general level of population health. MR (2006) also make the observation that using such a proxy may overstate the impact of HIV/AIDS as there is evidence that prevalence rates are higher among women than men in high prevalence countries and that the health status of a mother will have a greater impact on infant mortality. Empirical studies that have analysed the growth impact of HIV/AIDS have made use of both life expectancies and infant mortality rates as proxies for health capital.

3.2 Education as a complex input into the growth process

Recently, a number of researchers have moved beyond treating education as a homogenous concept to investigate how different levels of education impact on economic growth and development. Petrakis and Stamatakis (2002) investigate whether investment in the various levels of education contribute differently to the growth process. They find that different levels of education investment do, in fact, impact on growth differently and that the growth impact

varies depending on a country's stage of development. Specifically, primary and secondary education investments are more important for Less Developed Countries (LDCs) whereas higher education investment is more important for developed and advanced countries. As well as the potential for returns to investment in various levels of education to differ, Ramcharan (2004) demonstrates that the returns to one level of education investment depend on the stock of education capital available at other levels.

Conceptually, using a single measure of enrolment ratios as a proxy for education capital is problematic. A number of studies, such as that by MR (2006) and Tandon (2005), that use the secondary school enrolment ratio, potentially misses the growth impact of improved primary education enrolment rates for LDCs that have yet to approximate universal primary school attendance and also the growth effects of higher education for advanced countries that approximate universal secondary school attendance. Using a composite primary and secondary enrolment ratio is also problematic. When these two measures are aggregated in this way, the model cannot distinguish between an improvement in primary education enrolment and secondary enrolments, the growth effects of which may be different.

4 Model, Data and Methodology

The innovation of this study is to investigate the effect of HIV/AIDS on economic growth based on panel data regressions that extend the framework developed by MR (2006) by allowing a role for education capital accumulation and drawing on methodological advances in the panel data econometric literature by employing a system-GMM estimator representing an improvement over the traditional differenced-GMM estimator.[15]

The MR (2006) framework itself represents a development of the work of MRW (1992) and Knowles and Owen (1995).

The model developed herein differs from that of MR (2006) in two key ways:

1. Drawing on the recent literature regarding the potential role HIV/AIDS plays in the accumulation of education capital, in addition to the separate estimation of a health capital equation, a reduced form model of education capital is specified. Inter-linkages between health and education capital are explicitly modelled thus allowing for an investigation of whether HIV prevalence affects economic growth through both its role in the determination of health capital and education capital accumulation.

[15] See Section 4.3 for a detailed discussion.

2. The specification of educational capital has been broadened to include three indicators rather than one single measure to allow for the possibility that different levels of education capital accumulation can affect growth differently and that the growth impact varies depending on a country's stage of development.

Next the empirical growth model is presented. Following that, an explanation of the data is provided before going on to specify the growth, health capital and education capital equations to be estimated and the methodology employed.

4.1 The empirical growth model

The seminal empirical growth paper by MRW (1992) assumes identical Cobb-Douglas production functions for each country. Human capital, in the form of education, enters into the production function as an additional input. Islam (1995) extends this framework by incorporating a panel data approach. MRW (1992) assume that the growth rate of technology (g) and the capital depreciation rate (δ) are constant across countries. They allow the initial state of technology (A_0) to differ across countries. Technology is assumed to encompass not just technology in the typical sense, but also natural resource endowments and institutions, etc. Initial technology takes the form:

$$\ln (A_0) = a + \varepsilon$$

where 'a' represents a constant and 'ε' a country-specific shock term. MRW (1992) assume that the country-specific shock term is independent of population growth and investment in physical and human capital. This assumption is convenient in an econometric sense as it allows for the equation to be estimated simply, using OLS. Islam (1995) argues, however, that there is good reason to believe that the country-specific shock term may be correlated with the explanatory variables resulting in omitted variable bias. The panel data framework allows for this potential bias to be corrected. In addition to correcting for this potential bias, a panel framework provides many other well known additional benefits including: improved efficiency, more degrees of freedom, less risk of multicollinearity, and an ability to analyse the dynamics of adjustment (Baltagi, 2005).

Knowles and Owen (1995) extend the MRW (1992) framework to include both health and education capital as inputs into the production function in a cross-section framework. McDonald and Roberts (2002) incorporate health capital into a panel data framework. The model developed here represents an extension of these frameworks.

Here, the shortcomings of using a single indicator for education capital are addressed by specifying education capital in three forms; the skills and knowledge obtained through primary, secondary, and tertiary education.

A Cobb-Douglas production function with constant returns to scale and diminishing returns to each factor is given by:

$$Y_{it} = K_{it}^{\alpha} H_{it}^{\beta} P_{it}^{\psi} S_{it}^{\gamma} U_{it}^{\theta} (A_{it} L_{it})^{1-\alpha-\beta-\psi-\gamma-\theta} \quad (1)$$

$$\alpha, \beta, \psi, \gamma, \theta \in [0,1] \text{ and } \alpha + \beta + \psi + \gamma + \theta \in [0,1]$$

Where Y represents output, K is physical capital, A is the level of technology, L is labour, H is health capital and P, S and U are three forms of education capital – primary, secondary and tertiary respectively. The subscript *t* refers to time and *i* to country, $\alpha, \beta, \psi, \gamma,$ and θ are partial elasticities of output with respect to physical capital, health capital and primary, secondary and tertiary education capital respectively. It is assumed that an exogenously given fraction of output is saved and invested (w^K) in physical capital in each period and that the capital stock (both physical and human) depreciates over time at an exogenously given rate δ. It is also assumed that labour and technology grow at exogenously given rates *n* and *g* respectively.

It can be shown that the steady-state income per capita (y_{it}^*), in log form, may be expressed as:[16]

$$\ln y_{it}^* = Z \ln A_{i0} + g_t t - \left(\frac{Z\alpha}{1-\alpha}\right) \ln\left(n_i + g_t + \delta\right) + \left(\frac{Z\alpha}{1-\alpha}\right) \ln w_i^K + \left(\frac{Z\beta}{1-\alpha}\right) \ln h_{it}^*$$

$$+ \left(\frac{Z\psi}{1-\alpha}\right) \ln p_{it}^* + \left(\frac{Z\gamma}{1-\alpha}\right) \ln s_{it}^* + \left(\frac{Z\theta}{1-\alpha}\right) \ln u_{it}^* - (Z-1) \ln y_{i0}^* \quad (2)$$

where $h^*, p^*, s^*,$ and u^* denote per capita steady-state levels of health, primary, secondary and tertiary education capital respectively and y_{i0} represents the initial output per capita. The speed of convergence to the steady-state (λ), following Romer (2001, p.24), is given by:

$$\lambda = \left(n_i + g_t + \delta\right)\left(1 - \alpha - \beta - \psi - \gamma - \theta\right)$$

and $Z = \left(1 - e^{-\lambda t}\right)$.

[16] The derivation of this equation is not presented in detail because similar expressions have been derived in detail elsewhere (see for example, MRW (1992) and McDonald and Roberts (2002)). The only difference here is the inclusion of five indicators of human capital.

Equation (2) can be represented as a dynamic panel data model by using the conventional notation of the panel data literature as follows:

$$\ln y_{it}^{*} = \varpi \ln y_{i0}^{*} + \sum_{j=1}^{6} \Phi_j x_{it}^{j} + \eta_t + \mu_i + v_{it} \qquad (3)$$

where

$Z = \left(1 - e^{\lambda t}\right)$	$\Phi_5 = \left(\frac{Z\gamma}{1-\alpha}\right)$
$\varpi = 1 - Z$	$\Phi_6 = \left(\frac{Z\theta}{1-\alpha}\right)$
$\eta_t = g_t t$	$x_{it}^{1} = \ln\left(n_i + g_t + \delta\right)$
$\mu_i = ZA_{i0}$	$x_{it}^{2} = \ln w_i^{K}$
$\Phi_1 = -\left(\frac{Z\alpha}{1-\alpha}\right)$	$x_{it}^{3} = \ln h_{it}^{*}$
$\Phi_2 = -\Phi_1$	$x_{it}^{4} = \ln p_{it}^{*}$
$\Phi_3 = \left(\frac{Z\beta}{1-\alpha}\right)$	$x_{it}^{5} = \ln s_{it}^{*}$
$\Phi_4 = \left(\frac{Z\psi}{1-\alpha}\right)$	$x_{it}^{6} = \ln u_{it}^{*}$

and v_{it} represents the standard residual term.

Equation (3) represents a dynamic two-way fixed effects model.

4.2 The Data

This study uses data from a variety of sources, with the principal source being the World Development Indicators (WDI) database where various demographic, health and education indicators are drawn.[17] Education indicators are also drawn from databases of the UNESCO Institute of Statistics (2006). GDP and investment data are drawn from Penn World Tables (PWT) version 6.2 (see Heston et al., 2006). Proxies for nutritional status are drawn from the United Nations Food and Agriculture Organization database (FAO; 2006, 1998). Measures of the population at risk of malaria are obtained from data compiled by Gallup *et al.* (2001) and HIV/AIDS prevalence rates are obtained from UNAIDS (2006). Data on the quality of

[17] See World Bank (2006).

governance are drawn from International Country Risk Guide (ICRG) produced by the Political Risk Service Group (2006).

Table 2 lists all variables, their acronyms (in parentheses) and their sources. Appendix B describes the variables used in more detail.

Table 2: Variable Definitions and Sources

Variable	Source
Real GDP per capita (YC)	Penn World Tables 6.2 (2006)
Investment share of GDP (INV)	Penn World Tables 6.2 (2006)
Primary school enrolment ratio (PEN)	World Development Indicators (WDI) (World Bank 2006); UNESCO (2006)
Secondary school enrolment ratio (SEN)	WDI (World Bank 2006); UNESCO (2006)
Tertiary enrolment ratio (TEN)	WDI (World Bank 2006); UNESCO (2006)
School repetition rate (REP)	WDI (World Bank 2006); UNESCO (2006)
Student teacher ratio at the primary level (ST)	WDI (World Bank 2006); UNESCO (2006)
Government expenditure on education as a proportion of GDP (ED$)	WDI (World Bank 2006); UNESCO (2006)
Proportion of population aged 15 or under (POP15)	WDI (World Bank 2006)
Proportion of population living in urban areas (URB)	WDI (World Bank 2006)
Population growth rate (n)	WDI (World Bank 2006)
Life expectancy at birth (LE)	WDI (World Bank 2006)
Infant mortality rate (INF)	WDI (World Bank 2006)
Proportion of the population at risk of malaria (MAL)	Gallup *et. al* (1999)
Calorie intake per capita (CPC)	FAO (2006, 1998)
Adult prevalence of HIV (HIV)	UNAIDS (2006)
Index of governance (GOV)	Political Risk Service Group (2006)
Technological growth rate (g)	N/A
Capital depreciation rate (δ)	N/A

The data-set is grouped into panels of five-year periods from 1960 through 2004, although a number of the variables are only available from 1970 onwards. All the variables used are simple averaged data for each five-year period, with the exception of the malaria and investment in physical capital data (see Appendix B). The use of five-year averages is

common in the literature and is employed for various reasons: (i) to iron-out short-term volatility, (ii) to overcome measurement error that is often present in annual observations, and (iii) because some of the variables chosen are not available on an annual basis.

The available sample consists of 142 countries, of which are 32 are 'high-income' (HIC), 69 'developing and transitional', 41 'least developed countries', and 44 Sub-Saharan African. This sample contains within it all of the countries included in the studies by Tandon (2005) and MR (2006). There are nine 5-year time periods, i.e., 1960-1964, 1965-1969 etc. Since observations are not available for all countries and time periods for every variable, the sample represents an unbalanced panel. Appendix B describes, in more detail, the process via which the sample was chosen and contains a list of countries included. Tables 3 present summary statistics of the sample data.

4.3 The estimated equations

The GDP growth equation

The growth equation to be estimated is given by

$$\ln YC_{it} = \varpi \ln YC_{i,t-1} + \Phi_1 \ln\left(n + g + \delta\right)_{it} + \Phi_2 \ln INV_i + \Phi_3 \ln LE_{it} + \Phi_4 \ln PEN_{it} + \Phi_5 \ln SEN_{it}$$

$$+ \Phi_6 \ln TEN_{it} + \eta_t^1 + \mu_i^1 + v_{it}^1 \qquad (4)$$

Where the dependent variable, $\ln YC_{it}$, is the logarithm (to base e) of income per capita. Ideally, income per worker would have been used, but this series has not been updated in the Penn World Tables since 1990. Some comfort is taken from the fact that in a number of past studies involving estimation of Solow style growth models, results have, surprisingly, not been sensitive to the choice of per capita or per worker measures.[18]

The term, $\ln YC_{i,t-1}$, the one-period lagged log income per capita, is commonly known as the convergence term. The augmented Solow model predicts 'conditional convergence'. Conditional convergence means that countries with the same steady-state determinants (i.e. investment in physical and human capital, population growth, and technological growth) are expected to convergence to the same long-run steady-state level of income per capita, with

[18] See, for example, Hoeffler (2002).

Table 3: Summary Descriptive Statistics – Full and Sub-Samples

	2000-2004					1985-1989					1970-1974				
Variable[1]	**Obs**	**Mean**	**St. Dev.**	**Min**	**Max**	**Obs**	**Mean**	**St. Dev.**	**Min**	**Max**	**Obs**	**Mean**	**St. Dev.**	**Min**	**Max**
Full Sample (*i*=142)															
YC	142	9864	10358	388	50335	132	5510	5592	295	22409	130	1867	2136	72	12985
INV	142	13.7	7.6	1.8	34.5	132	14.8	8.3	2.6	47.4	130	16.9	10.4	1.5	59.6
ED$	108	4.6	1.9	0.6	9.6	121	4.1	2.0	0.4	10.9	104	3.84	1.57	0.60	8.50
GOV	120	17.8	5.8	5.2	29.6	115	15.6	6.5	4.8	28.9	0				
N	143	1.55	1.02	-0.92	3.78	143	2.08	1.34	-0.71	7.73	143	2.34	1.69	-1.24	15.78
POP15	142	32.7	10.4	14.2	50.4	143	36.8	9.8	17.1	49.7	142	39.6	8.2	20.4	49.4
URB	142	54.5	24.7	9.2	100	143	49.0	25.0	5.2	100.0	143	41.6	24.9	2.7	100.0
HIV	137	2.47	5.32	0.00	31.03	105	0.39	1.01	0.00	5.91	143	0.00	0.00	0.00	0.00
LE	142	64.6	13.1	37.3	81.5	143	62.6	10.8	39.3	78.3	141	57.0	11.5	35.1	74.8
MAL	135	0.38	0.45	0.00	1.00	135	0.44	0.44	0.00	1.00	135	0.51	0.42	0.00	1.00
CPC	127	2806	752	1261	4818	127	2614	569	1408	3835	132	2712	543	1515	3769
PEN	137	99.8	17.5	36.7	146.1	137	93.7	24.3	13.6	172.9	128	79.9	32.8	3.0	165.2
SEN	136	68.5	34.2	6.4	153.6	138	52.4	31.5	3.9	117.9	125	31.9	26.8	0.6	101.9
TEN	130	24.7	22.8	0.4	85.7	136	13.2	13.2	0.13	77.7	111	6.4	8.3	0.0	53.1
REP	117	0.09	0.09	0.00	0.40	100	0.11	0.10	0.00	0.41	63	0.13	0.10	0.00	0.34
ST	130	29.2	14.8	10.0	70.0	132	30.4	13.1	10.0	69.8	128	34.5	11.4	16.0	69.0
Sub-Saharan Africa (*i*=44)															
YC	44	2621	3478	388	17010	42	1537	1590	341	6529	42	586	572	157	3193
INV	44	8.5	5.5	3.3	30.7	42	8.9	4.6	2.6	18.9	42	11.0	8.5	1.7	42.4
ED$	30	4.0	1.8	0.6	9.6	33	3.4	1.7	0.4	6.8	25	3.4	1.2	1.1	5.8
GOV	32	13.3	3.8	5.2	19.8	31	12.7	4.4	5.1	21.8	0				
N	44	2.20	0.83	0.03	3.41	44	2.86	0.94	-0.71	4.97	44	2.51	0.96	-1.24	4.81
POP15	44	43.7	4.2	25.2	50.4	44	45.4	3.1	30.8	49.7	43	44.5	2.8	34.1	49.4
URB	44	34.9	15.0	9.2	81.5	44	27.9	12.6	5.2	65.1	44	19.5	10.8	2.7	47.9
HIV	44	6.85	7.90	0.02	31.03	44	0.87	1.45	0.00	5.91	44	0.00	0.00	0.00	0.00
LE	44	47.5	7.2	37.3	72.1	44	50.3	6.9	39.3	68.3	44	45.3	6.1	35.1	62.7
MAL	43	0.82	0.33	0.00	1.00	43	0.87	0.27	0.00	1.00	43	0.89	0.27	0.00	1.00
CPC	42	2167	477	1261	3419	42	2129	302	1408	2988	42	2218	337	1515	2963
PEN	41	89.5	23.2	37.3	134.9	40	78.4	35.1	13.6	172.9	41	55.8	33.5	8.0	165.2
SEN	40	30.9	18.7	6.4	87.6	40	21.9	16.3	3.9	68.2	41	8.2	6.0	1.1	30.7
TEN	39	3.7	3.6	0.4	15.5	40	2.2	2.0	0.1	10.8	33	0.7	0.8	0.0	4.1
REP	37	0.17	0.09	0.03	0.40	33	0.20	0.11	0.00	0.41	19	0.19	0.11	0.00	0.34
ST	40	45.8	11.8	24.3	70.0	39	42.8	13.5	19.0	69.8	39	43.9	11.4	24.0	69.0

Table 3: continued

	2000-2004					1985-1989					1970-1974				
Developing and Least Developed Countries (*i*=110)															
YC	109	4626	3899	388	17010	99	2551	2024	295	10769	97	887	714	72	3667
INV	109	11.5	6.2	2.3	30.7	99	11.9	6.2	2.6	34.0	97	13.7	8.3	1.7	42.4
ED$	81	4.29	1.89	0.61	9.59	88	3.78	2.12	0.40	10.90	75	3.57	1.40	0.60	7.70
GOV	87	15.2	3.9	5.2	24.9	82	12.7	4.4	4.8	21.8	0				
N	110	1.66	0.98	-0.92	3.41	110	2.27	1.09	-0.71	4.97	110	2.41	1.03	-1.24	6.14
POP15	110	36.5	8.6	14.9	50.4	110	40.7	6.9	21.0	49.7	109	42.5	6.0	20.4	49.4
URB	110	46.0	21.7	9.2	92.0	110	39.8	20.9	5.2	87.9	110	32.1	19.3	2.7	82.8
HIV	107	3.17	5.93	0.00	31.03	101	0.41	1.04	0.00	5.91	110	0.00	0.00	0.00	0.00
LE	110	60.3	12.3	37.3	78.3	110	58.8	9.7	39.3	74.6	108	53.0	10.0	35.1	71.1
MAL	110	0.50	0.45	0.00	1.00	110	0.57	0.41	0.00	1.00	110	0.63	0.39	0.00	1.00
CPC	101	2634	754	1261	4819	101	2438	500	1409	3835	106	2531	447	1515	3562
PEN	105	99.3	19.6	36.7	146.1	105	91.2	27.2	13.6	173.0	99	74.3	33.8	5.6	165.2
SEN	105	56.1	29.1	6.4	105.6	106	40.7	26.8	3.9	114.6	98	20.7	18.4	0.6	83.4
TEN	100	16.3	16.5	0.4	66.4	103	9.0	9.3	0.1	38.2	83	3.6	4.2	0.0	21.0
REP	94	0.11	0.09	0.00	0.40	76	0.14	0.11	0.00	0.41	47	0.16	0.09	0.00	0.34
ST	100	33.8	14.1	10.3	70.0	103	34.1	12.6	12.8	69.8	99	37.4	10.9	17.0	69.0
High Income Countries (*i*=32)															
YC	32	26225	6555	15256	50335	32	13868	3761	6112	22409	32	4625	2406	891	12985
INV	32	21.3	6.5	4.7	34.5	32	23.2	7.7	5.4	47.4	32	25.9	10.7	1.5	59.6
ED$	28	5.4	1.5	1.6	8.4	32	4.9	1.3	2.1	8.1	30	4.45	1.77	1.40	8.50
GOV	32	24.4	4.4	15.2	29.6	32	22.3	6.0	7.5	28.9	0				
N	32	1.13	1.01	0.14	3.78	32	1.41	1.80	-0.05	7.73	32	2.08	2.95	0.05	15.78
POP15	32	20.7	5.3	14.2	38.7	32	24.1	6.4	17.1	42.0	32	30.0	7.2	20.8	45.1
URB	32	80.6	12.5	55.7	100.0	32	77.1	13.3	46.3	100.0	32	71.3	15.8	39.6	100.0
HIV	32	0.29	0.51	0.00	2.93	6	0.23	0.56	0.00	1.37	32	0.00	0.00	0.00	0.00
LE	32	77.9	2.4	69.6	81.5	32	74.3	2.8	65.8	78.3	32	69.8	4.6	53.1	74.8
MAL	30	0.00	0.00	0.00	0.00	30	0.04	0.13	0.00	0.52	30	0.15	0.30	0.00	1.00
CPC	28	3386	356	2557	4079	28	3205	344	2198	3773	28	3353	323	2681	3769
PEN	31	101.6	8.3	68.1	120.5	32	101.6	9.1	68.8	128.8	30	100.0	14.4	45.3	122.8
SEN	31	105.6	20.0	70.7	153.6	32	87.2	15.9	41.8	117.9	30	64.6	21.0	12.1	101.9
TEN	30	50.3	20.9	11.2	85.7	32	25.8	15.4	2.6	77.7	29	14.1	11.3	1.4	53.1
REP	23	0.01	0.02	0.00	0.05	23	0.04	0.05	0.00	0.16	18	0.05	0.08	0.00	0.26
ST	30	15.2	4.3	10.0	31.5	30	18.6	6.2	10.0	37.0	29	26.4	8.5	16.0	57.0

Source: See Table 2

Notes: A full description of variables is provided in Appendix B

those countries starting at a lower levels of income per capita predicted to grow faster. The coefficient on $\ln YC_{i,t-1}$ is therefore expected to be positive and significant.

The variable $\ln(n+g+\delta)_{it}$ is often referred to as the 'capital widening' term. It is equal to the natural log of the sum of the population growth rate, technological growth rate and the capital depreciation rate. Following a convention in the literature, the sum of technological growth and the depreciation rate is here assumed to be uniform at 5%. Higher population growth is expected to reduce income per capita as it results in a dilution of capital. This is because, although an increase in population, absent any change in technology or capital, will raise total income, income per capita will fall in the presence of diminishing returns.

The variable $\ln INV_i$ represents the natural log of the ratio of domestic investment in physical capital to GDP. The investment rate is modelled as being exogenously given and time-invariant, and as such, for each country, the average value for the entire period from 1970 to 2004 is used.

The next four terms represent components of the stock of human capital. The stock of health capital is proxied by the natural log of life expectancy at birth, $\ln LE_{it}$. The knowledge and skills embodied in the population resulting from primary, secondary and tertiary education are proxied by the natural log of the gross enrolment ratio at each level of education: $\ln PEN_{it}$, $\ln SEN_{it}$ and $\ln TEN_{it}$ respectively. Each component of human capital is treated as endogenous and instrumented using their predicted values from separate equations that are detailed below. The coefficients on the physical capital investment term and the human capital stock terms are all expected to be positive, because an increase in all forms of capital is expected, *ceteris paribus*, to raise the productivity of labour and its returns.

The term η_t^1 represents a time-specific fixed effect and reflects time-specific improvements in technology that are assumed to be the same for all countries. These time-specific fixed effects are captured by including a time dummy variable for all but one of the nine 5-year time periods covered by this study.[19]

Country-specific effects are captured by allowing for variations in initial states of technology, represented by μ_i^1.

[19] Time-specific fixed effects are captured in the same way in the health and education equations.

Finally, the residuals are given by v_{it}^1.

The health capital equation

$$\ln LE_{it} = \Phi_7 \ln LE_{i,t-1} + \Phi_8 \ln YC_{i,t-1} + \Phi_9 URB_{it} + \Phi_{10} MAL_{it} + \Phi_{11} HIV_{it} + \Phi_{12} \ln PEN_{it}$$

$$+ \Phi_{13} \ln SEN_{it} + \Phi_{14} \ln TEN_{it} + \Phi_{15} \ln CPC_{it} + \eta_t^2 + \mu_i^2 + v_{it}^2 \qquad (5)$$

The equation presented here represents a population health production function. Health can be thought of as being 'produced' with a number of economic, social and environmental inputs. In the absence of a consensus in the literature on a structural model of the determinants of health capital at the aggregate level, a simple reduced-form health production function is estimated here.

The dependent variable, $\ln LE_{it}$, is the natural log of life expectancy at birth, a proxy for health capital. In the robustness testing in Section six, an alternative proxy – the infant mortality rate is used. Because the dependent variable is proxying for a stock, it is deemed necessary to include a one-period lagged dependent variable as one of the regressors.

Numerous past studies indicate that income levels are an important determinant of health status.[20] The expected sign on $\ln YC_{it}$ is positive as it is hypothesised that higher incomes allow for greater investment in one's health. Both current and one period lagged income per capita are considered.

Inclusion of a measure of public expenditures on health was considered, however cross-country comparable health expenditure data could not be obtained for a large enough time series and cross-section of countries.

The proportion of the population living in urban areas, URB_{it}, is included. A higher degree of urbanization may be associated with better access to health care services and information as well as the health services being more cost effective (Fayissa and Gutema, 2005). There is also the possibility, however, that a greater degree of urbanization may be associated with increased pollution and congestion and have a negative effect on health status (Fayissa and Gutema, 2005). A rapid increase in urbanization may also have a negative impact on health status if urban

[20] See, for example, Baldacci *et al.* (2004) and references within.

infrastructure such as sanitation is unable to keep pace. As such, the effect of urbanization on health may be positive or negative.

The proportion of the population at risk of malaria, MAL_{it} , as well as the adult HIV prevalence rate HIV_{it}, are included as two key indicators of a country's disease environment. The expected sign on both these variables is negative.

According to Grossman (1972), better educated individuals are more efficient producers of health. Education capital, specifically the natural log of primary and secondary education enrolment ratios, is therefore included in the health production function. Both current and one-period lagged values are considered. The expected sign on the education variables is positive. Tertiary education capital is included, but for different reasons. Secondary education is likely to provide a sufficient level of knowledge to appropriately address one's health needs, but tertiary education, by significantly raising the opportunity cost of poor health, may be expected to raise the incentive to invest in one's health.

Up to a particular threshold, better nutrition, proxied here by the natural log of calorie intake per capita ($\ln CPC_{it}$), leads to improved health and therefore the expected sign on this variable is positive.

Both country-specific, μ_i , and time-specific, η_t , fixed effects are included to allow for unobserved heterogeneity in population health across countries and time.

Finally, the residuals are given by v_{it}^2.

The education capital equations

$$\ln PEN_{it} = \Phi_{16} \ln PEN_{i,t-1} + \Phi_{17} \ln YC_{i,t-1} + \Phi_{18} ED\$_{it} + \Phi_{19} \left(ED\$_{it} * \ln GOV_{it} \right) + \Phi_{20} \ln LE_{it}$$
$$+ \Phi_{21} HIV_{it} + \Phi_{22} REP_{it} + \Phi_{23} POP15_{it} + \Phi_{24} URB_{it} + \eta_t^3 + \mu_i^3 + v_{it}^3 \qquad (6)$$

$$\ln SEN_{it} = \Phi_{25} \ln SEN_{i,t-1} + \Phi_{26} \ln PEN_{i,t-1} + \Phi_{27} \ln YC_{i,t-1} + \Phi_{28} ED\$_{it} + \Phi_{29} \left(ED\$_{it} * \ln GOV_{it} \right)$$
$$+ \Phi_{30} \ln LE_{it} + \Phi_{31} HIV_{it} + \Phi_{32} POP15_{it} + \Phi_{33} URB_{it} + \Phi_{34} REP_{it} + \eta_t^4 + \mu_i^4 + v_{it}^4 \qquad (7)$$

$$\ln TEN_{it} = \Phi_{35} \ln TEN_{i,t-1} + \Phi_{36} \ln SEN_{i,t-1} + \Phi_{37} \ln YC_{i,t-1} + \Phi_{38} ED\$_{it} + \Phi_{39} \left(ED\$_{it} * \ln GOV_{it} \right)$$

$$+\Phi_{40}\ln LE_{it}+\Phi_{41}HIV_{it}+\Phi_{42}URB_{it}+\eta_t^5+\mu_i^5+v_{it}^5 \quad (8)$$

Three reduced-form education capital production functions are estimated to reflect the skills and knowledge that are obtained via primary, secondary and tertiary education.

The dependent variable for the primary, secondary and tertiary education capital equations are the natural log of the primary enrolment ratio ($\ln PEN_{it}$), secondary enrolment ratio ($\ln SEN_{it}$) and tertiary enrolment ratio ($\ln TEN_{it}$) respectively. A defence for including what is conceptually a 'flow' measure of education as a proxy for its stock was provided in Section 3.1.

As in the health capital equation, income per capita is included as higher incomes are assumed to increase the resources available for individuals to invest in their own or their children's education. Additionally, the incentive to investment in education is greater in the presence of higher economy-wide average incomes due to the increased expected returns from the investment. Both the current and one-period lagged level of income per capita is considered for inclusion.

Following Baldacci *et al.* (2004), government expenditure on education as a proportion of GDP is included (ED$\$_{it}$). Greater government resources directed towards education is hypothesised to improve both access to and quality of schooling and hence participation. Additionally, the natural log of an index of the quality of a country's governance ($\ln GOV_{it}$) is included as an interaction variable with ED$\$_{it}$ to reflect the potential for the quality of governance to impact on the effect public spending has on education. Both the current and one-period lagged level of ED$ and the interaction term is considered for inclusion.

Health capital is included to reflect the possibility that it impacts on the incentive to invest in education, by changing its expected returns (Ferreira and Pessoa, 2003). Declining life expectancy is hypothesised to reduce investment in schooling and as such the coefficient is expected to be negative. Both the current and one-period lagged level is considered for inclusion

Adult HIV prevalence is included directly, as well as indirectly through the health capital variable, due to its potential impact on schooling quality. Haacker (2002a) highlights the potential strains on the schooling sector of high HIV prevalence among the teaching profession. These include; higher mortality among teachers leading to increased student-teacher ratios, a common indicator of schooling quality, and teacher deaths disrupting schooling in the event those teachers cannot be replaced immediately, a potential problem in rural areas with small schools.

The expected sign of the coefficient on this variable is negative as lower quality schooling might reduce the incentive to attend.

Following Baldacci *et al.* (2004), among others, the quality of education is included as proxied by either the repetition rate (REP_{it}) or student-teacher ratio (ST_{it}). The expected sign of the coefficient on this variable is negative as lower quality schooling might reduce the incentive to attend. These quality variables are not included in the tertiary education equation due to lack of data.

Mingat and Tan (2003, 1999) argue that a lighter demographic burden, here measured as the share of population below the age of 15 ($POP15_{it}$), reduces pressures on the education system and allows for more money to be spent per school-age child. The expected sign of the coefficient on this variable is therefore negative. This variable is not included in the tertiary education equation.

The proportion of the population in urban areas is included for similar reasons as for the health capital equation, that is, a higher degree of urbanization may be associated with easier access to education services as well as them being more cost effective.

In each equation, because the dependent variable is proxying for a stock, it is deemed necessary to include a one-period lagged dependent variable in each case. Additionally the secondary education equation includes a one-period lagged primary education capital term reflecting the fact that having a primary education usually a prerequisite for attending secondary school. Similarly, a one-period lagged secondary education term is included in the tertiary education equation.

Both country-specific, μ_i, and time-specific, η_t, fixed effects are included in each equation to allow for unobserved heterogeneity in education capital across countries and time.

Finally, the residuals are given by v_{it}^3, v_{it}^4, and v_{it}^5.

4.4 Estimation methodology

The estimation methodology employed addresses the problem of endogenous regressors in two respects. Firstly, in the growth equation, health and education capital are considered endogenous for reasons previously highlighted. This means that the health and education capital terms in equation (4) are potentially correlated with the residuals (v_{it}^1). Ordinary least squares (OLS) is biased and inconsistent as an estimator of equation (4). Secondly, in models with unobserved

country-specific fixed effects, μ_i, and the presence of lagged dependent variables, as in equations (4) through (8), OLS estimation is inconsistent because the lagged dependent variable will be correlated with the current period residuals producing upwardly biased coefficient estimates on the lagged dependent variable (Baltagi, 2005).

Here, two forms of Instrumental Variable (IV) estimation are used to solve the problem of endogenous regressors. The IV estimator, a method of moments estimator, represents an alternative principle of estimation to that of least squares. An instrument is defined as a variable that is correlated with the problematic explanatory variable but uncorrelated with the residuals. It can be shown that given the selection of such an instrument, the IV estimator is consistent (Hill *et al.*, 2001).

Two-stage least squares estimator

Potential endogeneity of health and education capital in equation (4) is addressed through the use of two-stage least squares regression (2SLS), a form of IV estimation. The 2SLS approach adopted here is the approach used by MR (2006) and Tandon (2005) in their analysis of the HIV/AIDS epidemic, except here, education capital, as well as health capital, is treated as endogenous.

In the first stage of 2SLS, potentially endogenous variables from the equation of interest are regressed on a set of valid instruments and also some or all of the other variables that appear in the equation of interest (the second-stage equation). The predicted values from these regressions are obtained. In the second stage, the equation is estimated as usual, except that the predicted values from the first stage regressions are used as 'instruments' for the potentially endogenous variables. These 'instruments' will be uncorrelated with the residuals and the slope estimators obtained in this second stage will be consistent (Hill *et al.*, 2001). Here, in the first stage, equations for health capital and primary, secondary and tertiary education capital are estimated and the predicted values used as instruments in the estimation of the growth equation.

Separate estimation of health and education capital production functions also allows for a theoretically sound approach to identifying the macroeconomic effects of the HIV/AIDS epidemic.

Generalised Method of Moments Estimator

The second econometric issue that needs to be addressed involves the dynamic nature of equations (4) through (8). As mentioned above, OLS is an inconsistent estimator of models with unobserved country-specific fixed effects, μ_i, and the presence of lagged dependent variables, even if the residuals do not exhibit any autocorrelation (Baltagi, 2005). One standard approach for panel models with unobserved country-specific effects is to use the 'within-groups estimator' or 'fixed-effects estimator' (FE estimator). This approach rests on the assumption of strict exogeneity of the explanatory variables conditional on μ_i (Wooldridge 2002, p.266). The approach involves taking deviations from the individual country means, known as the 'within transformation' and apply pooled OLS to the transformed equation. In the context of a simple model with one explanatory variable and without time-specific fixed effects given by

$$y_{it} = \beta_1 + \beta_2 x_{it} + \mu_i + \nu_{it}$$

calculating the means for each variable and subtracting gives

$$y_{it} - \bar{y} = (\beta_1 - \beta_1) + \beta_2 (x_{1it} - \bar{x}_{1i}) + (\mu_i - \bar{\mu}_i) + (\nu_{it} - \bar{\nu}_i)$$

$$y_{it} - \bar{y} = \beta_2 (x_{1it} - \bar{x}_{1i}) + \nu^*_{it}$$

If we apply OLS to this time demeaned equation it produces a consistent and unbiased estimator (Wooldridge, 2002).[21]

Unfortunately, the use of this FE estimator is not appropriate in the context of models including lagged dependent variables as regressors. Including a lagged dependent variable, $y_{i,t-1}$, in the simple model above, the within transformation produces

$$y_{it} - \bar{y} = \beta_2 (x_{1it} - \bar{x}_{1i}) + \beta_3 (y_{i,t-1} - \bar{y}) + \nu^*_{it}$$

It can be shown that $(y_{i,t-1} - \bar{y})$ will be negatively correlated with ν^*_{it}, even if ν_{it} does not exhibit autocorrelation (Baltagi, 2005). The coefficient on the lagged dependent variable will be downwardly biased when 't' is very large.

There are various IV methods to deal with the endogeneity of lagged dependent variables in panel models with fixed effects. One such widely used method, particularly suited to situations where

[21] In the context of equations (4) through (8) where time-specific fixed effects are also present, a full set of dummy variables can be added to control for this before applying the 'within transformation'.

the time dimension of the dataset is quite short, is proposed by Arellano and Bond (1991). This method, commonly known as difference-GMM, effectively takes first-differences of the equation to remove fixed effects and then uses appropriate instruments for the lagged differenced dependent variable. Implementation rests on two assumptions; that there is no autocorrelation in the residuals, v_{it}, and that initial conditions are predetermined, that is, y_{i1} is not correlated with the subsequent residual terms v_{it} (where $t \neq 1$). Using the Generalised Method of Moments framework of Hansen (1982), a set of moment conditions generated by the two assumptions can be exploited allowing two-period lagged levels and deeper as instruments in the first-differenced equation.[22] This will produce consistent estimates of the coefficient on the lagged dependent variable when the time dimension is fixed and $i \to \infty$.

MR (2006) and Tandon (2005) employ this difference-GMM estimator in their analysis of the HIV/AIDS epidemic. This is where the current analysis makes a significant methodological departure. This study is the first to apply the system-GMM approach, an approach initially proposed by Arellano and Bover (1995) and further developed by Blundell and Bond (1998), to the analysis of the cross-country growth effects of the HIV/AIDS epidemic.

In addition to the assumptions underlying the method of Arellano and Bond (1991), if we also assume that the initial changes in the dependent variable and any other endogenous regressors are uncorrelated with the fixed effect, μ_i, another set of moment conditions will become available. This additional moment condition allow for lagged first differences to be used as instruments for the equation in levels.[23] The approach is then to estimate, using GMM, a system of equations; a set of first-differenced equations with lagged levels used as instruments, and a set of equations in levels with lagged differences used as instruments. For predetermined, but endogenous variables, such as lagged dependent variables, the appropriate instrument set contains lags one and earlier of the instrumenting variable for the equation in differences, and lags zero and earlier of the instrumenting variable in differences for the levels equation. For current realisations of endogenous variables, the appropriate instrument set contains lags two and earlier of the instrumenting variable for the equation in differences, and lags one and earlier of the instrumenting variable in differences for the levels equation.

[22] Please refer to Arellano and Bond (1991) for a detailed treatment.
[23] Please refer to Arellano and Bover (1995) for a detailed treatment.

The calculation of this system-GMM estimator is discussed in detail in Blundell and Bond (1998). The estimates of the coefficients in the system can be shown to be consistent and these estimates are then used to recover the underlying structural parameters of interest.

The system-GMM has been shown to be superior to difference-GMM in certain circumstances, ones which are present in this study. Blundell and Bond (1998) show, through the use of Monte Carlo simulations, that in either of the two following circumstances lagged levels will be weak instruments for the first-differenced equation and that the difference-GMM estimator will exhibit downward finite sample bias especially when the time dimension is small: (i) The coefficient on the lagged dependent variable is close to unity, that is, the dependant variable is highly persistent, or (ii) the variance of the individual fixed effect, μ_i, increases relative to the variance of the residuals, ν_{it}.

The system-GMM estimator has been found to preform well relative to difference-GMM in the context of the estimation of production functions (Blundell and Bond, 2000) and particularly in the estimation of empirical growth models, where difference-GMM estimation was found to produce seriously biased estimates (Bond *et al.*, 2001).

The findings of Blundell and Bond (2000) and Bond *et al.* (2001) cast doubt on the accuracy of earlier estimates of the macroeconomic impact of HIV/AIDS obtained from growth regressions employing the difference-GMM estimator. Other practical benefits of system-GMM that are particularly useful in the current study include;

1. it is less sensitive to transient measurement error (Bond *et al.*, 2001), and
2. one less time period is lost in the estimation process relative to difference-GMM.

The validity of the instruments underlying system-GMM can be checked using a range of specification tests. The first such test that is valid for any GMM estimation of panel models, discussed in detail in Arellano and Bond (1991), is a test for autocorrelation in the first-differenced residuals ($\Delta\nu_{it}$). Mathematically, the process of first-differencing the data will induce first-order autocorrelation of the differenced residuals. The absence of second-order autocorrelation in the first-differenced equation is equivalent to there being no autocorrelation in the levels residuals (ν_{it}). The moment conditions underlying both difference-GMM and system-GMM estimates rely on the absence of autocorrelation in the levels residuals, ν_{it}, and hence the

Arellano and Bond (1991) test for second order autocorrelation in the first-differenced residuals are reported in various tables in Section 5.

A valid instrument is one that is independent of the residuals. In some circumstances it is possible to test for such independence. If we have more instruments than are needed to identify an equation, the well known Hansen (1982) J-test of overidentifying restrictions is available.[24] Another important specification test is the 'difference in Sargan' test, or the 'C' test. This test allows for an assessment of the validity of the additional instruments used in system-GMM relative to difference-GMM. The statistic underlying this test is the difference between two Hansen J-test statistics.[25] Both the J and C tests are reported in the results section.

In Section 5, only the results from system-GMM estimation are presented as these are *a-priori* considered superior to OLS estimation, FE estimation and difference-GMM estimation. Knowledge of the direction of the theoretical bias resulting from these alternative estimation techniques is useful, because confidence in the consistency of the system-GMM estimator can be gained if it is observed to lie somewhere between the OLS and FE estimates and above the difference-GMM estimates (Bond, 2002). Results from these alternative estimation techniques are therefore presented as part of the robustness checking in Section 6.

5 Results

The health and education equations as described by equations (5) through (8) are initially expanded by adding a set of squared terms for each variable, and each variable interacted with the income per capita and HIV prevalence term (except for the lagged dependent variable). Variables that were individually insignificant at the 10% level, and where inclusion added little explanatory value to the model, were removed. The final model specifications are reported in Tables 4 through 10. No quadratic or interaction terms were added to the growth equation (4), nor were insignificant variables dropped, due to the desire to retain the equation in a form that reflects its strong theoretically underpinnings.[26] For both difference-GMM and system-GMM estimation, one-step and two-step options are available.[27] For each equation one-step GMM estimates are reported as opposed to two-step estimates. Although in the case of system-GMM, two-step

[24] See Arellano and Bond (1991) for a discussion of this test in the context of GMM estimation of dynamic panels.
[25] See Arellano and Bover (1995) for a discussion of this test in the context of system-GMM estimation.
[26] An exception to this is the omission of some, but not all, of the various education capital terms.
[27] See Baltagi (2001) page 134.

estimators are theoretically more efficient, Monte Carlo simulations have highlighted two problems with two-step estimators, including;

1. the estimator converges to its asymptotic distribution relatively slowly and can be biased in finite samples, and
2. the usual asymptotically valid standard errors associated with the estimator exhibit downward bias in finite samples (Bond *et al.*, 2001).

Although a solution to the second problem has been proposed by Windmeijer (2005), in the form of a finite sample correction to the downwardly biased standard errors, the first problem remains, and, in any case, Monte Carlo simulations show that the efficiency gains of two-step estimators are often only small (Bond *et al.*, 2001).

For each equation, time-specific fixed effects, η_t, are captured by including a full set of time dummy variables. These time dummies are retained in the model if they are found to be jointly significant.

5.1 Results for the health equation

Table 4 presents the system-GMM estimates for equation (5). Results are reported for both the full sample and a sub-sample of developing countries. Sub-sample analysis is achieved by incorporating a full set of intercept and slope-shifting dummy variables into the equation. Incorporation of dummy variables in this way was considered preferable to separate estimation of the sub-sample of developing world countries as it improves the efficiency of the estimation by maximizing degrees of freedom.[28]

Income per capita, nutritional status and primary education capital, proxied by the primary enrollment ratio, all have the expected positive signs and are strongly significant for both the full and developing world sample.[29] Additionally, the coefficients on the proportion of the population at risk of Malaria (MAL), as well as adult HIV prevalence rate (HIV), are negative and strongly significant. The degree of urbanization was not found to be significant and was thus dropped from both estimated equations. It is possible that the insignificance of urbanization could be due

[28] This is also the approach used in the estimation of the education and the growth equations.
[29] Both current and lagged income were considered and one period lagged income provides marginally better results.

to the theorized competing positive and negative effects outlined in Section 4.3 offsetting one another.

Table 4: Results for the health equation (World and Developing world sample)

Dependent variable: ln(LE)	**Estimation technique System-GMM (one-step robust estimates)**	
	Full sample (a)	**Developing World sample (b)**
L.ln(LE)	0.5921 (0.000)***	0.6175 (0.000)***
L.ln(YC)	0.0275 (0.002)***	0.0255 (0.010)***
Ln(CPC)	0.0526 (0.007)***	0.0507 (0.012)**
L.ln(PEN)	0.0544 (0.001)***	0.0510 (0.001)***
MAL	-0.0294 (0.023)**	-0.0272 (0.032)**
HIV	-0.0131 (0.000)***	-0.0129 (0.000)***
Constant	0.8466 (0.000)***	0.7890 (0.001)***
Time dummies	Chi2[6]=27.0 (0.000)***	Chi2(6)=22.63 (0.001)***
Wald	Chi2[12]=11984 (0.000)***	Chi2(18)=37885 (0.000)***
Hansen (J-test)	Chi2[26]=38.7 (0.053)*	Chi2(44)=48.95 (0.281)
Difference-in-Sargan (C-test)	Chi2[6]=11.3 (0.081)*	Chi(12)=9.42 (0.667)
AR(2)	Z=-1.70 (0.090)*	Z=-1.69 (0.091)*
Instrument count	39	63
Observations	678	678
Countries	125	125

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(LE) is instrumented for with system-GMM style instruments.

Secondary education capital, proxied by the secondary enrollment ratio, was not found to be significant in either estimated equation. It is possible that this reflects the likelihood that primary education is a sufficient level of education to equip individuals with the knowledge to appropriately address their health needs. Tertiary education capital, proxied by the tertiary enrolment rate, was also omitted as it was found to be insignificant.

For the full world sample, the marginal effect of a full percentage point increase in HIV prevalence on health capital, proxied by life expectancy at birth, calculated at the world mean life

expectancy, is a loss of 2.1 years. For the developing world, calculated at the developing world mean life expectancy, the marginal effect is a loss of 1.9 years.[30]

The Arellano and Bond (1991) test statistics for no second-order autocorrelation in the estimated residuals in first-differences are asymptotically distributed as standard normal variables under the null hypothesis of no autocorrelation. For both the full and developing world, we would fail to reject the null at the 5% level of significance. The Hansen test, based on a null hypothesis that the instruments are independent of the estimated residuals is based on a J-statistic that has a Chi-squared distribution and is reported in Table 4. Also reported is the 'difference-in-Sargan' C-test of the validity of the additional instruments used in the system-GMM estimator to that of the difference-GMM estimator. The C-test statistic is based upon the difference between the Hansen J-test statistics from the model estimated with the system-GMM estimator and the model estimated with the difference-GMM estimator. This resulting C-statistic follows a Chi-Squared distribution, with failure to reject the null hypothesis confirming the validity of the sub-set of additional instruments used in system-GMM relative to difference-GMM. We fail to reject the null for the both the Hansen and difference-in-Sargan tests for both the full and developing world sample.

Overall, in both estimated equations, the specification tests validate the instrument set underlying the lagged dependent variable. The specification tests underlying the full sample does raise some concerns as the null hypothesis for each of these tests would have been rejected based on a 10% significance level. Nonetheless, the specification is retained for now, but has been subjected to robustness testing in Section 6.

5.2 Results for the education capital equations

The results from the primary education equation for the full world sample indicate that only lagged primary education capital and the fixed time-effects are significant. In the developing world sample, only lagged income per capita is found to be significant at the 5% level (*p*-value of 0.049), along with the time dummy variables. This outcome most likely reflects the fact that for a large part of the sample period, all but the world's poorest countries have achieved close to universal primary school enrolments. In any case, primary education capital is found to be highly insignificant in the growth equation for both the full world and developing world sample and therefore the results for the primary education equation are not presented here. This result is

[30] These calculations ignore feedback effects of lower life expectancy on income.

consistent with the findings of Petrakis and Stamatakis (2002) who show that the growth impact of different levels of education depends on a country's stage of development and that primary education is only likely to be an important source of growth for the least developed nations. Consistent with this finding, in the sub-sample analysis of Sub-Saharan African countries in Section 5.4, primary education capital is found to be relatively more important for growth than in the full world and developing world sub-sample.

Moving onto secondary education capital, Table 5 presents the results for the full world and developing world sub-sample. Both income per capita and government spending on education as a proportion of GDP (ED$) are strongly statistically significant with the expected positive sign.[31] The magnitude of the effects of the two variables on secondary education capital is greater in the developing world sample. This is an intuitive result, because in many HICs, differences in income per capita or ED$ are unlikely to explain changes in secondary education enrolments. This is because for a significant portion of the time frame under consideration, almost universal secondary education enrolment has been achieved. ED$ interacted with an index of governance was found to be insignificant and therefore dropped from the model.

Lagged primary education capital exhibits the expected positive sign and is statistically significant, reflecting that obtaining primary education is a prerequisite for achieving secondary education.

Health capital has the expected positive impact on secondary education capital and is highly significant for both the full and developing world sample. The proportion of a country's population under 15 (POP15) has a negative sign and is statistically significant, reflecting that a lighter demographic burden, by reducing pressure on the ability of governments to maintain a properly resourced education system, can lead to increased participation. The magnitude of the coefficient is more than twice as large in the developing world sample as in the full sample.[32] Neither HIV nor urbanization were found to be statistically significant in either the full and developing world sample and were therefore dropped from the model.

[31] For income, ED$, and health capital, both current and lagged levels were considered. Current values were found to provide marginally better results in each case.

[32] Although not reported here, the sub-sample analysis shows that amongst a sample of HICs, POP15 is not significant.

Table 5: Results for the secondary education equation (World and Developing world sample)

Dependent variable: ln(SEN)	Estimation technique System-GMM (one-step robust estimates)	
	Full sample (a)	Developing World sample (b)
L.ln(SEN)	0.6369 (0.000)***	0.4843 (0.000)***
ln(YC)	0.0529 (0.025)**	0.0982 (0.020)**
L.ln(PEN)	0.2029 (0.007)***	0.3096 (0.001)***
ED$	0.0201 (0.000)***	0.0290 (0.000)***
ln(LE)	0.4104 (0.001)***	0.5164 (0.003)***
POP15	-0.0045 (0.027)**	-0.0121 (0.001)***
Constant	-1.4942 (0.004)***	-2.0280 (0.004)***
Time dummies	Chi2[6]=44.71 (0.000)***	Chi2[6]=30.60 (0.000)***
Wald	Chi2[12]=7396 (0.000)***	Chi2[19]=7807 (0.000)***
Hansen (J-test)	Chi2[26]=54.96 (0.001)***	Chi2[52]=65.30 (0.102)
Difference-in-Sargan (C-test)	Chi2[6]=9.15 (0.165)	Chi2[12]=10.74 (0.551)
AR(2)	Z=-2.63 (0.008)***	Z=-2.82 (0.005)***
Instrument count	39	72
Observations	746	746
Countries	137	137

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(SEN) is instrumented for with system-GMM style instruments. Instruments are restricted to lags two and earlier of the instrumenting variable for the equation in differences, and lags one and earlier of the instrumenting variable in differences for the levels equation.

Turning to the specification tests, for the full sample, the null hypothesis of the AR(2) test and J-test is strongly rejected. The natural response to the presence of autocorrelation was to include an additional lag of the dependent variable. This did not, however, solve the problem. An alternative strategy is to estimate the equation regardless of the presence of autocorrelation, but alter the instrument set. For system-GMM, for predetermined endogenous regressors such as lagged dependent variables, given the presence of AR(2) in the differenced residuals, we must restrict the instrument set to lags two and earlier of the instrumenting variable for the equation in differences, and lags one and earlier of the instrumenting variable in differences for the levels equation, as long as there is no third-order autocorrelation in the differenced residuals. The results for the full-sample, as reported in Table 5, are those based on the restricted instrument set

as described above.[33] Unfortunately, the null hypothesis from the Hansen J-test is still strongly rejected, suggesting the instruments remain poor in any case.

For the developing world, the instrument set is restricted in the same way due to the presence of AR(2) in the differenced residuals. In this case, the null hypothesis underlying both the J-test and C-test validate the instruments used.

Notwithstanding the failure of the Hansen J-test for the full-sample, the specification is retained for now, however the equation is subjected to scrutiny in the robustness testing in Section 6.

Moving onto the tertiary education equation, the coefficient on the lagged income per capita term is positive, as expected, and highly statistically significant. Lagged secondary education capital has a positive and statistically significant effect, reflecting that secondary education is often a prerequisite for obtaining a tertiary education. Health capital, proxied by life expectancy at birth, is found to have a positive effect on tertiary education capital and this is statistically significant at the 5% level, consistent with the proposition that rising life expectancy increases the incentive to invest in tertiary education, due to increasing the expected returns from the investment.[34]

Unlike in the health and secondary education equation, the sign of urbanization is positive and statistically. A higher degree of urbanization has two theorised effects; (i) increased urbanization corresponds with easier access to education services for a larger proportion of the population, and (ii) it enables governments to provide tertiary education institutions more cost effectively. Its significance in the tertiary education equation only, could reflect the relative priority for governments of tertiary education services versus secondary education and health services. In developing countries, provision of secondary education institutions and health services would likely have a higher priority than the provision of tertiary education institutions. Given this assertion, governments of less urbanized developing countries may simply not fund tertiary education institutions as it would prove too costly to provide them in a way such that they are easily accessible by a large proportion of the population. The magnitude of the coefficient for urbanization suggests it is more important in the sub-sample of developing countries, consistent with this argument.

HIV is found to have a negative and significant effect on tertiary education capital, whilst it is found to be insignificant in the secondary education equation. A negative effect was theorised to reflect supply-side pressures on the education sector whereby high HIV prevalence among

[33] No third-order autocorrelation in the differenced residuals was detected.

[34] Estimation was performed with both current and lagged income and health capital. The best results were achieved by including one period lagged income but current health capital.

teaching staff leads to lower quality and quantity of education services. One explanation for the relevance of HIV in this equation but not in the secondary education equation, may be that in developing countries with severe epidemics, government finances are being diverted into the health sector to tackle the epidemic and that these additional resources may come at a sacrifice to funding for the tertiary education sector (amongst other areas), as opposed to a less favourable alternative of sacrificing funding to secondary and primary education.

Table 6: Results for the tertiary education equation (World and Developing world sample)

Dependent variable: ln(TEN)	**Estimation technique GMM-Sys (one-step robust estimates)**	
	Full sample	**Developing World sample**
L.ln(TEN)	0.3720 (0.000)***	0.3533 (0.000)***
L.ln(SEN)	0.3310 (0.000)***	0.3278 (0.000)***
L.ln(YC)	0.1428 (0.013)**	0.1491 (0.025)**
ln(LE)	0.6895 (0.021)**	0.6631 (0.026)**
URB	0.0079 (0.002)***	0.0120 (0.000)***
HIV	-0.0157 (0.007)***	-0.0139 (0.017)**
Constant	-1.1335 (0.274)	-1.1481 (0.275)
Time dummies	Chi2[6]=42.00 (0.000)***	Chi2[6]=35.62 (0.000)***
Wald	Chi2[11]=4696 (0.000)***	Chi2[18]=10297 (0.000)***
Hansen (J-test)	Chi2[30]=39.44 (0.116)	Chi2[53]=50.67 (0.565)
Difference-in-Sargan (C-test)	Chi2[6]=3.72 (0.714)	Chi2[12]=5.70 (0.930)
AR(2)	Z=-1.91 (0.056)*	Z=-1.94 (0.052)*
Instrument count	43	72
Observations	644	644
Countries	134	134

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(TEN) is instrumented for with system-GMM style instruments.

The null hypothesis underlying the AR(2) test, J-test and C-test cannot be rejected for either the full or developing world sample, supporting the validity of the instrument set. The null hypothesis underlying the AR(2) test only marginally fails to reject at the 5% level of significance and therefore the sensitivity of the results to restricting the instrument set is investigated in the robustness Section 6.

5.3 Results for the growth equation

Table 7 presents the system-GMM estimates from equation (4) with predicted values from the health and education equations used as instruments for the endogenous health and education regressors.

Table 7: Results for the growth equation (World and Developing world sample)

Dependent variable: $\ln(YC)_t$	**Estimation technique GMM-Sys (one-step robust estimates)**		
	Full sample		**Developing World sample**
	(a)	**(b)**	**(c)**
Convergence coefficient	0.8793 (0.000)***	0.8535 (0.000)***	0.8133 (0.000)***
ln(INV)	.0386 (0.022)**	.0398 (0.035)**	-.0006 (0.981)
$\ln(n + g + \delta)$	-.1576 (0.000)***	-.1978 (0.000)***	-.1540 (0.004)***
ln(SEN)	.05860 (0.072)*	.0274 (0.243)	.1109 (0.004)***
ln(TEN)		0.0359 (0.151)	
ln(LE)	.3139 (0.045)**	.3142 (0.023)**	.3214 (0.014)**
Constant	-.9260 (0.019)**	-.9572 (0.014)**	-.5533 (0.165)
Time dummies	Chi2[6]=84.09 (0.000)***	Chi2[6]=65.35 (0.000)***	Chi2[6]=85.23 (0.000)***
Wald	Chi2[12]=49505 (0.000)***	Chi2[13]=39401 (0.000)***	Chi2[18]=70942 (0.000)***
Hansen (J-test)	Chi2[38]=31.93 (0.745)	Chi2[38]=30.71 (0.794)	Chi2[73]=75.69 (0.392)
Difference-in-Sargan (C-test)	Chi2[13]=10.07 (0.688)	Chi2[13]=17.09 (0.195)	Chi2[26]=22.03 (0.687)
AR(2)	Z =-0.01 (0.988)	Z =0.01 (0.989)	z =0.39 (0.697)
Instrument count	51	52	92
Observations	812	774	812
Countries	136	136	136

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(YC) and L2.ln(YC) are instrumented for with system-GMM style instruments. ln(LE), ln(SEN) and ln(TEN) are all instrumented for with their corresponding predicted values from the first stage regressions.

Primary education capital did not improve the explanatory power of the model and was not included in the final specification. An additional convergence term, $\ln(YC)_{t-2}$, is included in the final model specification. This was because the estimated equation including only $\ln(YC)_{t-1}$

suffered from serious autocorrelation, as evidenced by a strong rejection of the AR(2) test. In Table 7, the 'convergence' coefficient therefore represents the sum of the coefficients on $\ln(YC)_{t-1}$ and $\ln(YC)_{t-2}$ and the reported *p*-value is based on a test of joint significance.

The results for the growth equation are generally very pleasing. The joint significance of the two lagged income per capita terms, for both the full and developing world sample, confirms the presence of conditional convergence. For the full sample, the results from two alternative estimated equations are reported, one including secondary education capital (estimated equation (a)) and the other including both secondary and tertiary education capital (estimated equation (b)). Under both specifications (a) and (b), the sign on the investment term is positive, as expected, and significant at the 5% level. The coefficient for 'capital widening' ($\ln(n + g + \delta)$) has the expected negative sign and is highly significant under both specifications. This is in contrast to the findings of MR (2006) and Tandon (2005) who find in their full and sub-sample analysis either an insignificant effect or an unexpected positive effect.

Health capital is found to have a positive and significant effect on income per capita at the 5% level in both specifications (a) and (b). The coefficient on the secondary education capital term exhibits the expected positive sign and, although not significant at the conventional 5% level, the *p*-value of 0.072 would lead us to reject the null under a less conservative 10% significance level. Although not reported in Table 7, when excluding secondary education capital and including tertiary education capital, the coefficient for tertiary education capital exhibits the expected positive sign with a reported *p*-value of 0.079. In estimated equation (b), the signs of both secondary and tertiary education capital are positive as expected. Although the individual *p*-values are 0.243 and 0.151 respectively, a test of joint significance of the two education capital terms would lead to a rejection of the null at a 10% level. It is also reassuring to note that when moving from specification (a) to (b), the estimated coefficients in the model remain quite stable. Although we cannot conclude that the effects of either secondary or tertiary education capital on income per capita, or both together, are significant at the 5% level, observing the expected positive sign, along with *p*-values of less than 0.10, represents a promising result relative to many other empirical growth studies.

Turning to the developing world sample, the capital widening term has the expected negative coefficient and is statistically significant.[35] The coefficient on the investment term is extremely

[35] Although not reported here, the impact of the capital widening term for a sub-sample of HICsis not significantly different from zero.

close to zero and highly insignificant. Health capital is found to have the expected positive effect on income per capita and this is significant at the 5% level.

As in the case of the full sample, the coefficient on secondary education capital is positive, but the magnitude of the coefficient is clearly larger and it is also significant at the 1% level. Although not reported in Table 7, when including tertiary education capital instead of secondary education capital, the effect of tertiary education capital on income per capita is found to be insignificantly different from zero. When including both education terms, the effect of secondary education capital remains positive and significant at the 5% level, whilst again the effect of tertiary education capital is found to not be significantly different from zero.

These results are consistent with a number of common observations from the development literature. Firstly, in relation to the relative importance of investment in physical and human capital in developed versus developing countries, there is a commonly held view that, in developing countries, relative to developed countries, investment in human capital is comparatively more important than investment in physical capital. The differing magnitudes of the coefficients on investment and the education capital terms in the full and developing world sample are consistent with this view.[36] Although we might expect a smaller effect of investment in physical capital in the developing world sub-sample, the observed zero effect does contradict expectations. Additionally, a comparison of the results between the full and developing-world sub-sample are consistent with the findings of Petrakis and Stamatakis (2002) and others that the more advanced a countries stage of development, the greater is the importance of higher levels of education for economic growth and the lesser is the importance of lower levels of education. Finally, for both the full and developing world equations, the validity of the instrument set for the lagged dependent variables cannot be rejected.

5.4 The case of Sub-Saharan Africa

Sub-Saharan Africa has, and continues to experience, the highest HIV prevalence rates in the world. As reported in Table 1 in Appendix A, the weighted average prevalence rate in the region in 2005 was 6.1, meaning that 6.1% of all 15-49 year olds are HIV positive. This rate of prevalence is nearly four times greater than the Caribbean, the second worst affected region. It is

[36] In further support of this commonly held view, the impact of investment in physical capital for HICs has a magnitude more than twice that found in the full-sample analysis and is highly significant.

considered worthwhile to separately investigate the economic impact of HIV prevalence amongst Sub-Saharan African countries, reflecting that many of these countries are experiencing a scale of epidemic very different to the majority of the rest of the world.

Once again, sub-sample analysis is achieved through the use of dummy variables with the reported coefficients incorporating the slope effects captured by the interaction dummies.[37]

Table 8: Results for the health equation (Sub-Saharan Africa sub-sample)

Dependent variable: ln(LE)	**Estimation technique System-GMM (one-step robust estimates)**
	Sub-Saharan Africa
L.ln(LE)	0.6703 (0.000)***
L.ln(YC)	0.0226 (0.004)***
ln(CPC)	0.0513 (0.003)***
L.ln(PEN)	0.0331 (0.002)***
HIV	-0.0112 (0.000)***
Constant	0.6326 (0.000)***
Time dummies	Chi2[6]=32.85 (0.000)***
Wald	Chi2[15]=13744 (0.000)***
Hansen (J-test)	Chi2[52]=69.45 (0.053)*
Difference-in-Sargan (C-test)	Chi2[12]=14.80 (0.253)
AR(2)	Z=-1.54 (0.125)
Instrument count	68
Observations	680
Countries	126

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(LE) is instrumented for with system-GMM style instruments.

The results for the health equation are presented in Table 8. The same variables appear as significant as those in the estimated equation for the full and developing world sample, with the exception malaria, which is found to be insignificant and thus dropped from the model. Based on the reported specification tests, the validity of the instrument set for the lagged dependent variable cannot be rejected at the 5% level.

For the primary education equation, reported in the left-hand column of Table 9, lagged health capital and ED$ have a positive and strongly significant impact. The effect of income per capita

[37] Only when the interaction terms produced coefficients that were very small in magnitude and highly statistically significant were the dummy variable terms dropped from the final model specification.

on primary education capital is positive, as expected, but is only significant at the 10% level. These results contrast with that of the full sample and developing world sample, where only the lagged primary enrolment rates, time dummy variables, and, in the case of the developing world sample, income per capita, were found to be significant.

Table 9: Results for the primary and secondary education equations (Sub-Saharan Africa sub-sample)

Sub-Saharan Africa	**Estimation technique GMM-Sys (one-step robust estimates)**	
	Dependent variable: ln(PEN)	**Dependent variable: ln(SEN)**
L.ln(PEN)	0.7266 (0.000)***	0.2085 (0.005)***
L.ln(SEN)		0.5368 (0.000)***
ln(YC)	0.0409 (0.087)*	0.0895 (0.090)*
L.ln(LE)	0.0305 (0.023)**	0.1215 (0.024)**
ED$	0.0297 (0.005)***	0.0578 (0.000)***
POP15		-0.0083 (0.001)***
URB		0.0055 (0.074)*
Constant	0.7236 (0.021)**	-0.5347 (0.350)
Time dummies	Chi2[6]=23.46 (0.000)***	Chi2[6]=29.48 (0.000)***
Wald	Chi2[13]=1917 (0.000)***	Chi2[18]=7112 (0.000)***
Hansen (J-test)	Chi2[30]=34.34 (0.267)	Chi2[26]=50.94 (0.002)***
Difference-in-Sargan (C-test)	Chi2[6]=6.41 (0.379)	Chi2[6]=13.40 (0.034)**
AR(2)	Z=-0.61 (0.541)	Z=-2.90 (0.004)***
Instrument count	44	44
Observations	757	746
Countries	137	137

Notes:

7. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
8. L. denotes lagged value.
9. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
10. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
11. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
12. L.ln(PEN) and L.ln(SEN) are instrumented for with system-GMM style instruments. For the secondary education equation, instruments are restricted to lags two and earlier of the instrumenting variable for the equation in differences, and lags one and earlier of the instrumenting variable in differences for the levels equation.

The particularly strong relevance of ED$, highlights the importance for those countries that have yet to achieve the important goal of universal primary school attendance to devote resources to the education sector as a means to achieve that goal. The significance of health capital, proxied by life expectancy, in the primary education equation may be reflecting that for those Sub-Saharan African countries with particularly low life expectancies, young children are staying at

home to care for ill parents due to the absence of social safety nets in the form of care or financial assistance for those in ill-health. The specification tests clearly validate the instrument set underlying the lagged dependent variable.

Turning to the secondary education equation, both income per capita and ED$, have the expected positive. ED$ is found to have a strongly significant effect, whilst income per capita is only significant at the 10% level. As for the full and developing world sample, ED$ interacted with an index of governance was not found to be significant and therefore dropped from the model. HIV was also found to be insignificant and dropped from the model. Lagged primary education capital and lagged health capital are found to have the expected positive sign and are statistically significant. The sign of POP15 is negative, as expected, and is highly statistically significant. Unlike in the full and developing world equations, the degree of urbanization is included and has the expected positive sign, albeit with significance only at the 10% level. This is likely to be reflecting that in very poor countries with a large rural population, governments may struggle to provide accessible secondary education institutions or, alternatively, that parents force children to stay at home to assist with maintenance of the family properties.

Notable other differences between the Sub-Saharan African sub-sample and the full and developing world sample is the larger magnitude of the coefficient for ED$ and the smaller magnitude of the coefficient for health capital. The results of the AR(2) test confirm the presence of second-order autocorrelation in the differenced residuals and therefore, as for the full and developing world sample, the instrument set underlying lagged secondary education capital is restricted. Unfortunately, the null hypothesis underlying the Hansen test is still strongly rejected, suggesting the instruments remain poor in any case. The specification is retained for now, but is subjected to robustness testing in Section 6. With regards to the tertiary education capital equation, results are not reported here as tertiary education capital is subsequently found to be highly insignificant when included in the growth equation.

Three alternative specifications are investigated for the growth equation. In specification (a), both primary and secondary education capital is included, whilst specifications (b) and (c) include primary education capital only and secondary education capital only respectively. In each specification there is evidence of conditional convergence and that the speed of convergence is slower, as indicated by the greater magnitude of the coefficient on the lagged dependent variables, than in the full and developing world sample. This observed slower convergence rate is

to be expected because the variations in income per capita are relatively smaller amongst Sub-Saharan African countries.

Table 10: Results for the growth equation (Sub-Saharan Africa sub-sample)

Dependent variable: ln(YC)	**Estimation technique GMM-Sys (one-step robust estimates)**		
	Sub-Saharan Africa		
	(a)	**(b)**	**(c)**
Convergence coefficient	0.9282 (0.000)***	0.9239 (0.000)***	0.9436 (0.000)***
ln(INV)	-0.0010 (0.001)***	0.0047 (0.001)***	-0.0067 (0.001)***
ln(n + g + δ)	-0.2028 (0.001)***	-0.2288 (0.000)***	-0.2013 (0.000)***
ln(PEN)	0.0329 (0.209)	0.0513 (0.261)	
ln(SEN)	0.0225 (0.625)		0.0254 (0.635)
ln(LE)	0.2507 (0.073)*	0.3368 (0.089)*	0.2766 (0.020)**
Constant	-1.0856 (0.007)***	-1.4778 (0.008)***	-1.1478 (0.003)***
Time dummies	Chi2[6]=96.42 (0.000)***	Chi2[6]=105.3 (0.000)***	Chi2[6]=101.21 (0.000)***
Wald	Chi2[20]=82943 (0.000)***	Chi2[18]=71847 (0.000)***	Chi2[18]=84128 (0.000)***
Hansen (J-test)	Chi2[75]=71.40 (0.596)	Chi2[75]=72.98 (0.544)	Chi2[75]=70.60 (0.622)
Difference-in-Sargan (C-test)	Chi2[26]=23.88 (0.583)	Chi2[26]=22.68 (0.651)	Chi2[26]=22.24 (0.676)
AR(2)	Z=-0.20 (0.845)	Z=-0.57 (0.572)	Z=-0.28 (0.780)
Instrument count	96	94	94
Observations	810	817	827
Countries	136	136	138

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. AR(2) denotes the Arellano and Bond (1991) test of autocorrelation of the first-differenced residuals of order two.
4. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
5. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis. The p-value corresponding to the time dummies is for a null hypothesis of all time dummies being jointly zero. The p-value corresponding to the 'Wald' is for a null hypothesis of all the parameters in the model, excluding the constant, being jointly zero.
6. L.ln(YC) and L2.ln(YC) are instrumented for with system-GMM style instruments whilst ln(LE), ln(SEN) and ln(PEN) are all instrumented for with their corresponding predicted values from the first stage regressions.

The coefficient on the capital widening term has the expected negative sign and is highly statistically significant in each specification. Similar to the developing world sub-sample, investment in physical capital is found to have no effect on income per capita, contrary to expectations. Health capital exhibits the expected positive sign in each specification, however the precision of this estimate is sensitive to the way in which education capital is incorporated. When primary education capital is included individually or jointly with secondary education capital, health capital is found to be significant only at the 10% level. When only secondary education capital is included, or when education capital is excluded entirely, health capital is significant at

the 5% level. Tertiary education capital is highly insignificant and not included in any of the reported specifications. In all cases, the impact of primary or secondary education capital, either individually or jointly, is found to be insignificantly different from zero. The coefficient is, however, observed to have the expected positive sign, which is at least an improvement on many empirical growth studies that find when including both education and health capital, education capital often exhibits the opposite sign to what would be expected. Results are suggestive that primary education capital is at least as important as secondary education capital for economic growth amongst Sub-Saharan African countries, consistent with a common finding in the development literature that, for less developed countries, investments in lower levels of education have greater economic growth effects than for more developed countries.[38]

Finally, under all specifications, the validity of the instruments set underlying the lagged dependent variables cannot be rejected, as evidenced by the results of the AR(2), Hansen, and 'difference-in-Sargan' tests.

5.5 The economic impact of HIV prevalence

The impact of HIV prevalence on income per capita can be derived from the coefficients of the estimated equations. The elasticities of income per capita with respect to adult HIV prevalence, calculated at the relevant sample means, are reported in Table 11, along with comparisons to the estimated elasticities reported in previous similar cross-country econometric studies.

For the full world sample, the total effect of a 1% rise in adult HIV prevalence is, on average, a decrease in income per capita of between 0.138% and 0.146%, more than twice as large the estimates of other similar studies. This corresponds to a semi-elasticity of between -13.78 and -14.13%, implying that a full 1 percentage point rise in the HIV prevalence rate will result in between 13.78% and 14.13% reduction in income per capita, *ceteris paribus.* These estimates are significantly larger than those found by MR (2006) and Tandon (2005) even if the role the epidemic is having on education capital accumulation is ignored.

For the developing world, the elasticity is estimated to be -0.125% and the semi-elasticity -10.66%. The estimated impact due solely to health capital accumulation is similar to that of MR (2006) but, taking into the role of education capital accumulation, much larger. Relative to the full world sample, we are more certain of the role education capital accumulation plays in the

[38] See, for example, Petrakis and Stamatakis (2002).

economic growth consequences of the epidemic, with secondary education capital's impact on economic growth found to significant at the 1% level in Table 5.

Table 11: Elasticity of income per capita with respect to adult HIV prevalence

Sample	Impact through health			Impact through education	Total impact		
		MR	Tandon			MR	Tandon
World[1]:		-0.050	-0.052			-0.050	-0.052
Specification A	**-0.104**			**-0.042**	**-0.146**		
Specification B	**-0.081**			**-0.057**	**-0.138**		
Developing World	**-0.076**	-0.082		**-0.049**	**-0.125**	-0.082	
Sub-Saharan Africa[2]:		-0.585				-0.585	
Specification A	**-0.386**			**0 to -0.124**	**-0.386 to -0.510**		
Specification B	**-0.535**			**0 to -0.106**	**-0.535 to -0.641**		

Notes:
MR denotes the study by McDonald and Roberts (2006) and Tandon denotes the study by Tandon (2005).
The breakdown of the impact between health and education is achieved by first calculating the full impact and then the health impact is determined by setting the coefficient on education capital in the growth equation to zero. The impact through education is calculated as the difference between the two.
[1] Specification A corresponds to the growth equation that includes secondary education capital. Specification B corresponds to the growth equation including secondary and tertiary education capital.
[2] Specification A corresponds to the growth equation that includes primary education capital. Specification B corresponds to the growth equation including primary and secondary education capital.

For Sub-Saharan Africa, incorporating the estimated coefficients on education capital in the growth equation, a 1% rise in HIV prevalence is estimated to lead to a fall in income per capita of between 0.51% and 0.64% on average. A full one percentage point increase in prevalence is estimated to lead to between 20.6% and 21.6% decrease in income in per capita. Given the *p*-values reported in Table 10, we cannot, at any conventional level of significance, conclude that investment in education has any impact on income per capita amongst the sub-sample of Sub-Saharan African countries. Given the lack of statistical significance, we can only be confident of an estimated impact of a 1% rise in HIV prevalence on income per capita of between -0.386% and -0.535%. A full one percentage point rise in HIV prevalence is estimated to lead to a decline of between 15.6% and 21.6% in income per capita. The upper end of these estimated range of impacts for the Sub-Saharan sample are similar to those estimated by MR (2006).

6 Robustness testing

Robustness of the results presented in Section 5 is investigated in the following ways:

1. Sensitivity of the results from using infant mortality rates as an alternative to life expectancy as a proxy for health capital are analysed.
2. Sensitivity of the results to using protein intake per capita as an alternative to calorie intake per capita as a proxy for nutritional status are analysed.
3. Sensitivity of the results to restricting the GMM instrument sets as a response to the possible presence of second-order autocorrelation in the differenced residuals are analysed.
4. Comparison of the results obtained using alternative estimators such as OLS, FE and difference-GMM estimator are made.

6.1 Sensitivity to alternative health proxies

Overall results were not found to change significantly when using infant mortality rates as an alternative proxy for health capital. Additionally, results were not sensitive to using protein intake per capita as an alternative indicator of nutritional status.

6.2 Sensitivity to restricting the GMM instrument set

In the health equation and the tertiary education equations for the full and developing world sample, the null hypothesis of the AR(2) test although not rejected at the 5% level, would have been rejected at the 10% level. If second order autocorrelation did indeed exist in the differenced residuals, but not third order, one should restrict the instrument set underlying the lagged dependent variable. The results from the equations estimated by restricting the instrument set to lags two and earlier of the instrumenting variable for the equation in differences and lags one and earlier of the instrumenting variable in differences for the levels equation, are presented in Table 12 in Appendix C, along with the estimates from Tables 4 and 6 for ease of comparison.[39] For the health equation, in both the full and developing world sample, estimates remain relatively unchanged to restricting the instrument set. The tertiary education equation for the full world sample, however, is sensitive to this change. Restricting the GMM instrument set for the tertiary education equation leads to the coefficient on lagged tertiary education capital rising considerably, whilst coefficients on the other variables fall. Lagged secondary education capital

[39] The tertiary education equation is only shown for the full world sample because tertiary education capital does not have a significant impact on income per capita in the developing world and Sub-Saharan African samples.

becomes insignificant and income per capita and life expectancy are now only significant at the 10% level. The coefficient on HIV prevalence remains highly significant with only a slight fall in magnitude. In Table 11, the elasticity of income per capita with respect to HIV prevalence for the full world, based on the growth equation specification that includes tertiary education capital, was estimated to be -0.138%. Recalculating using the alternative point estimates from Table 12 produces an estimated elasticity of -0.156%. If we strictly apply the 5% significance level and assume that the effect of changes in income per capita and life expectancy on tertiary education capital is zero, the estimated elasticity falls to -0.116%.

6.3 Comparison to alternative estimators

In this exercise we do not seek to observe insensitivity of the results to the use of alternative estimators, but in fact, wish to see a wide range of results. In Section 4.4 it was stated that system-GMM estimation was considered *a-priori* superior to alternative estimation techniques. When estimating dynamic panels, the coefficient on the lagged dependent variable will be upwardly biased under OLS. When the time dimension is small, the coefficient on the lagged dependent variable will be downwardly biased when applying the FE estimator. The coefficient on the lagged dependent variable will be downwardly biased when using the difference-GMM estimator if the time dimension is small and the dependent variable is highly persistent (Blundell and Bond, 1998). According to Bond (2002), confidence in the consistency of the system-GMM estimator can be gained if estimates are observed to lie somewhere between the OLS and FE estimates and above the difference-GMM estimates.

Confidence in the consistency of the results presented in Section 5 can therefore be assessed by estimating each equation using OLS, FE and difference-GMM. A summary of the outcome of this exercise is presented in Table 13 and more detailed results are provided in Table 14 in Appendix C.

In large part, the results suggest we can be confident in our system-GMM estimates. The only situations where the ideal pattern of estimates was not observed were in the health equations and the growth equation for Sub-Saharan Africa. In the health equations, although system-GMM estimates of the lagged dependent variable were found to lie below OLS and above difference-GMM estimates, they were observed to be lower than those obtained from FE estimation. In the growth equation for Sub-Saharan Africa, the system-GMM estimates were observed to lie above the theoretically upwardly biased OLS estimates.

Table 13: Estimated coefficient on the endogenous lagged dependent variable under system-GMM estimation observed to lie …

	Below OLS	Above FE	Above difference-GMM
Health equations	✓		✓
Secondary education equations	✓	✓	✓
Tertiary education equation (Full sample)[1]	✓	✓	✓
Primary education equation (Sub-Saharan Africa)	✓	✓	✓
Growth equations (Full and developing world sample)[2]	✓	✓	✓
Growth equation (Sub-Saharan Africa sample)[3]		✓	✓

Notes

1. In the tertiary equation, the system-GMM estimate of the coefficient on the lagged dependent variable is observed to lie above the one-step difference-GMM estimate but not the two-step alternative.
2. For the full sample, it includes the specification including only secondary education capital as well as the specification including both secondary and tertiary education capital.
3. Includes the specification including only primary education as well as the specification including both secondary and tertiary education capital.

The pattern of results observed for the health equation warrants further investigation. The first response to the doubts raised regarding the consistency of the system-GMM estimator for the health equation was to consider changing the specification by considering the inclusion of both the one and two period lagged dependent variable, a strategy that was pursued for the growth equations. Pursuing this alternative specification resulted in very little change in the final calculated elasticity of income per capita to HIV prevalence, and, although fixed effects estimation continued to produce estimates under this alternative specification that were observed to lie above the system-GMM estimates, the size of the gap was not as large as previously.

Next, we investigate how the estimated elasticities of income per capita with respect to HIV prevalence would change if the health capital equations are estimated using FE, notwithstanding that in theory the coefficient on the lagged dependent variable is known to be downwardly biased. For the full, developing world and Sub-Saharan African samples, the estimated coefficient on lagged health capital is somewhat higher than the system-GMM estimates (see Table 14). In each case, the magnitude of the coefficient on HIV falls slightly, but *p*-values close to zero are still observed. In each case, the magnitude of the coefficient on the lagged income per capita term falls and is estimated with less precision. The overall impact of this change in estimation strategy on the estimated elasticity of income per capita with respect to HIV prevalence is reported in Table 15, along with the estimated elasticities arising from system-GMM.

Table 15: Elasticity of income per capita with respect to adult HIV prevalence under alternative estimation strategies for the health capital equation

	System-GMM	FE
Full world	-0.138 to -0.146	-0.174 to -0.184
Developing world	-0.125	-0.160
Sub-Saharan Africa	-0.386 to -0.641	-0.444 to -0.518

For the full and developing world sample, FE estimation of the health capital equation results in a material increase in the calculated elasticity of income per capita with respect to HIV prevalence. For Sub-Saharan Africa, the resulting elasticities remain within the range of results obtained under system-GMM estimation of the health capital equation.

In summary, results are quite robust and we have a considerable amount of confidence in the use of the system-GMM estimator. There are some exceptions to this broad conclusion. Firstly, the tertiary education equation is found to be sensitive to restricting the instrument set. The resulting changes in the magnitude and precision of the estimates lead to a small decline in the estimated elasticity of income per capita to HIV prevalence. Secondly, there are some doubts regarding the consistency of the system-GMM estimator for the health capital equations. Further investigation suggests that this is most likely leading to an understatement of the economic impact of the HIV/AIDS epidemic. These two concerns do not alter the broad conclusions that arise from the analysis in this paper; being that the economic impacts are large, and considerably larger than those estimated in previous similar studies. In fact, for the full and developing world sample, the combined effect of the sensitivity of the tertiary education equation and the doubts regarding the consistency of the GMM estimator underlying the health equations, is that we are more likely to be slightly underestimating the economic impact of the HIV/AIDS epidemic than overestimating. Lastly there are some doubts regarding the consistency of the system-GMM estimator for the Sub-Saharan Africa growth equation. Further investigation did not identify any obvious solutions to this problem.

7 Summary and Conclusions

The objective of this paper was to examine empirically the impact of HIV prevalence on economic growth. The analysis was motivated by an emerging literature that considers the growth impact of the HIV/AIDS epidemic due to its potential to slow human capital accumulation through education. Based on a sample of 142 countries over 45 years, a system of

equations for economic growth, health capital, and education capital are estimated. The economic growth model represents a variation of the augmented Solow model that includes human capital in the form of both health capital and multiple indicators of education capital. Both forms of human capital are treated as endogenous and inter-linkages between them are explicitly captured. The innovations of the analysis are twofold: (i) it represents the first econometric assessment that pays particular attention to human capital accumulation through education within an augmented neoclassical growth framework, and (ii) it implements dynamic panel system-GMM estimator that is know to be superior to the more traditional difference-GMM estimator under particular circumstances.

Increased adult HIV prevalence is found to reduce life expectancy, a proxy for health capital, which in turn slows economic growth. Reduced economic growth then reduces the resources available for future investment in health and education. This in turn feeds back into economic growth process. Additionally, an HIV/AIDS induced decline in average life expectancy has a direct and statistically significant impact on investment in education.

For the full world sample, the estimated impact of a 1% rise in adult HIV prevalence is a reduction in income per capita of between -0.116% and -0.156%, an impact that is substantially higher than that identified in previous cross-country empirical research. In a sub-sample of developing countries, the estimated impact is -0.125%. A material component of this estimated impact, particularly in the developing world sample, is through the epidemics capacity to slow education capital accumulation. These results are considered to be robust. The least satisfactory results are for the Sub-Saharan African sub-sample where no statistically significant effect of investment in education on income per capita is detected. Additionally, robustness testing raised some doubts regarding the consistency of the GMM estimator underlying the Sub-Saharan growth equation, and therefore in turn, the reliability of the estimated impact of rising adult HIV prevalence on income per capita.

The actual economic damage resulting from the HIV/AIDS epidemic could in fact be even larger than that identified in this study as the analysis does not consider the potential for the epidemic to slow physical capital accumulation and technological growth – impact channels that are potentially important. A shortcoming of our analysis, one that would also suggest that the estimated impacts are conservative, is that the chosen proxy for education capital may not be capturing the detrimental impact HIV/AIDS could have on informal human capital accumulation, that is, the skills of the workforce obtained either through job experience or on-the-job training.

In addition, the proxies adopted in this study for education capital, although defendable, do have obvious shortcomings. There is currently a lot of focus in human capital literature on developing improved proxy measures, and therefore any subsequent research should draw upon such improved measures as they become available.

In conclusion, this study finds that the HIV/AIDS epidemic involves not only devastating human and social costs, but also large economic costs. The magnitude and sources of the estimated impacts are consistent with the broad conclusions arising from an emerging body of literature that identifies as an important impact channel the potential for HIV prevalence to reduce economic growth by undermining the accumulation of education capital. This literature had to date investigated this impact channel through simulations of theoretical overlapping generations or CGE models based on individual or small groups of African countries. Here, the impact of the HIV/AIDS epidemic on cross-country economic growth, incorporating this previously overlooked impact channel, was directly estimated using data for 142 countries spanning 45 years.

The broad policy implications that have arisen from this emerging literature are supported by our results. These policy implications include; (i) that in countries considered at particular risk of increased HIV prevalence in the future, from a purely economic cost-benefit perspective, devotion of substantial preventative resources is justified to avoid large long-run economic damage, and (ii) for those countries already in the midst of a severe epidemic, not only should resources be devoted to tackling the epidemics spread and improving the quality and productivity of life for HIV sufferers, they should also be directed to maintaining participation in, and quality of, the education system that in turn would prevent a breakdown in the transmission of human capital across generations.

Appendix A

Table 1: HIV prevalence in the 15-49 years age group as a percentage of all persons 15-49 years of age

	2005 %	2003 %
Africa (Worst Eight)		
Swaziland	33.4	32.4
Botswana	24.1	24.0
Lesotho	23.2	23.7
Zimbabwe	20.1	22.1
Namibia	19.6	19.5
South Africa	18.8	18.6
Zambia	17.0	16.9
Mozambique	16.1	16.0
Rest of the World (Worst Eight)		
Haiti	3.8	3.8
Bahamas	3.3	3.3
Trinidad & Tobago	2.6	2.6
Belize	2.5	2.1
Guyana	2.4	2.4
Suriname	1.9	1.7
Papua New Guinea	1.8	1.6
Cambodia	1.6	2.0
Weighted averages by region		
Sub-Saharan Africa	6.1	6.2
Caribbean	1.6	1.5
North America	0.8	0.7
Eastern Europe & Central Asia	0.8	0.6
South & South-East Asia	0.6	0.6
Latin America	0.5	0.5
Oceania	0.3	0.3
Western & Central Europe	0.3	0.3
North Africa & Middle East	0.2	0.2
East Asia	0.1	0.1

Source: UNAIDS (2006)

Appendix B

B1 Variable Description

This section provides a brief description of the variables considered in this study

Economic and governance indicators

- **Real GDP per capita (YC):** is equal to GDP adjusted for Purchasing Power Parity (PPP) divided by the average population during a year.[40] Adjusting for PPP results in a common set of prices in a common currency allowing for real quantity comparisons. These quantities are expressed in current prices. This variable was sourced from PWT 6.2 (2006).
- **Investment share of GDP (INV):** is equal to domestic investment (in physical capital) divided by GDP (in current prices) and was sourced from PWT 6.2 (2006). In the model's specification, investment is assumed exogenously given and constant over time for each country. As such, the average value for each country over the entire period 1970 to 2004 is used.
- **Index of governance (GOV):** This index was constructed using a data set on institutional quality indicators produced by the IRIS Centre of the University of Maryland from the International Country Risk Guide (ICRG) – a monthly publication of Political Risk Services (PRS). Five indicators are summed together to produce an overall index of governance. These indicators include indexes of corruption, law and order, democratic accountability, bureaucracy quality, and investment profile. The indices of corruption, law and order, and democratic accountability are measured on a scale of zero to six (where higher values represent low risk). Bureaucracy quality is measured on a scale of zero to four, whereas investment profile is measured on a scale of zero to twelve. Each index is standardized to a scale of zero to six before summing to produce an overall index of governance. Thus 30 represents the highest possible quality of governance for a country.
- **Capital depreciation rate (δ):** is a time invariant, exogenously given rate at which all forms of capital depreciate.
- **Technological growth rate (g):** is equal to the annual improvement in Total Factor Productivity (TFP) and is assumed to be exogenously given. Economists define TFP as the actual measured growth of output minus the growth rate expected from increases in all forms of capital and labour.

Health indicators

[40] See Heston *et al*. (2006) for a detailed description of all data series obtained from PWT 6.2.

- **Life expectancy at birth (LE):** is equal to the number of years a newborn would live if prevailing patterns of mortality were to stay in place throughout their life. This variable was sourced from the WDI (2006) database.
- **Infant mortality rate (INF):** Proportion of infant deaths (deaths of individuals under one year of age) per 1000 live births. This variable was sourced from the WDI (2006) database.
- **Calorie intake per capita (CPC):** is equal to the total amount of calories from all food sources consumed by the population during the year. It is expressed on a per person, per day basis. This variable is derived from national food balance sheets produced by FAO (Food and Agricultural Organisation of the United Nations). Food available for consumption is estimated as a balancing item by considering total national supply and deducting the amounts used for feed, seed, industrial uses and waste. The new online database released in 2006 utilises an improved methodology to compute food consumption. FAO have restated the time-series back to 1990 utilising this new methodology and these figures were used for the post-1990 period. Pre-1990 data was sourced from the FAO CD-ROM (1998). For the period from 1990-1998 estimates based on the old methodology were observed to occasionally be materially different to that under the new methodology. As such, pre-1990 data was generated by taking the 1990 estimate from the new time series and applying the percentage changes based on the old series to generate a series dating back to 1970.
- **Adult prevalence of HIV (HIV):** is equal to the estimated number of people aged 15–49 with HIV divided by the total population of 15-49 year olds. The time series estimates of HIV prevalence are sourced from UNAIDS (2006) and represent an improvement on previously available data due to better estimation methods and improved data availability. UNAIDS have only made available to the general public the 2005 and 2003 estimates based on this new methodology. UNAIDS were kind enough to provide a full time series back to 1980 using the new methodology for use in this study on the understanding that individual country time series estimates were not published. The time series provided did not include some Western European and North American high income countries and as such only the observation in the publicly available UNAIDS (2006) report are used.
- **Proportion of the population at risk of Malaria (MAL):** is equal to the percentage of the population living in areas of high malaria risk in a country in a particular year. The figures are based on estimates of the world distribution of malaria and a 1995 estimate of

the world population distribution. This variable was sourced from Gallup *et al.* (1999). Data is only available at four points in time - 1946, 1966, 1982 and 1994. Between 1966 and 1982 straight-line interpolation is applied to produce estimates for the 1970-74 and 1975-79 time intervals. The 1982 observation is used as the 1980-84 observation. Similarly, straight-line interpolation is applied between 1982 and 1994. The observation for the 2000-04 period is produced by extending the ten-year trend to 1994 forward. If this process resulted in a number greater than one or less than zero, then it was changed to one or zero respectively. The 1995-1999 observation was generated by applying straight-line interpolation between the 1990-1994 and 2000-2004 figures.

Education indicators

In 1997 there was a change in the International Standard Classification of Education (ISCED) and this new classification was applied from 1998 onwards. The United Nations Educational, Scientific and Cultural Organization (UNESCO) warn that comparisons between post-1997 and pre-1997 data must be made with caution. Data for primary, secondary & tertiary enrolment ratios, student-teacher ratios, repetition rates and government expenditure on education as a proportion of GDP, are all affected. The WDI database only contains data on these variables from 1998 onwards. Pre-1998 data is available from the UNESCO databases.

- **Primary school enrolment ratio (PEN):** is equal to the total (gross) enrolments at primary school level, regardless of age, divided by the population of the age group that typically corresponds to primary education. Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.
- **Secondary school enrolment ratio (SEN):** is equal to the total (gross) enrolments at secondary school level, regardless of age, divided by the population of the age group that typically corresponds to secondary education. Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.
- **Tertiary enrolment ratio (TEN):** is equal to the total (gross) enrolments at tertiary level, regardless of age, divided by the population of the age group that typically corresponds to tertiary education. In equations (4), (5) and (8) TEN enters in natural log form. For some developing countries, tertiary enrolments rates are very close to zero and the log function is

poorly behaved near zero. Therefore, TEN is first multiplied by 100 before applying the natural log. Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.

- **Student teacher ratio at primary level (ST):** is equal to the average number of pupils per teacher at the primary level of education in a given school year. Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.
- **School repetition rate (REP**): is equal to the proportion of pupils enrolled in a given grade at a given school year who study in the same grade in the following school year. The rate used here is the average of the repetition rates for each grade at the primary school level for the primary education equation (6) and the average of the repetition rates for each grade at the secondary school level for the secondary education equation (7). Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.
- **Government expenditure on education as a proportion of GDP (ED$):** is equal to total public expenditure on education (current and capital expenses) divided by GDP for a given year. Data from 1998 onwards was sourced from the WDI database, whilst pre-1998 data was sourced from the UNESCO databases.

Demographic indicators

- **Population growth rate (n):** is equal to the annual percentage change in population and is sourced from the WDI database.
- **Proportion of population aged 15 or under (POP15):** is equal to the population aged 15 or under divided by the total population and is sourced from the WDI database.
- **Proportion of population living in urban areas (URB):** is equal to the population living in urban areas, divided by the total population and is sourced from the WDI database

B2 The Sample

The WDI database contains data for 208 countries or territories and represents the widest coverage of all the data sources utilized**.** Inclusion in this database does not necessary mean that data is available for each of the indicators for each time period**.** Using the coverage in this

database as a starting point, countries were progressively removed from the sample if they met one of two conditions:

1. The country did not exist in its current form for the whole of the time period (for example, members of the former Soviet Union, Germany, Yugoslavia, East Timor etc.)
2. Data was missing entirely for the majority of variables or for the majority of time periods for a majority of variables

This process resulted in the inclusion of 147 countries. A further five small oil producing nations were removed from the sample. This was due to the well-known unreliability of income per capita measures for these countries from PWT 6.2. The resulting final sample consists of 142 countries. The countries included are; Algeria, Angola, Argentina, Armenia, Australia, Austria, The Bahamas, Bahrain, Bangladesh, Barbados, Belgium, Belize, Benin, Bhutan, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of Congo, Republic of Congo, Costa Rica, Côte d'Ivoire, Cuba, Cyprus, Denmark, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Finland, France, Gabon, The Gambia, Ghana, Greece, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Islamic Republic of Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea Republic, Kuwait, Lao People's Democratic Republic, Latvia, Lebanon, Lesotho, Liberia, Libya, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Rwanda, Samoa, Saudi Arabia, Senegal, Sierra Leone, Singapore, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syrian Arab Republic, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Uganda, United Kingdom, United States, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Zambia, and Zimbabwe.

Appendix C

Table 12: Sensitivity of the health and tertiary education equations to restricting the instrument set.

Dependent variable: ln(LE)	**Estimation technique System-GMM (one-step robust estimates)**			
	Full sample		**Developing World sample**	
	Standard GMM instrument set	**Restricted GMM instrument set**	**Standard GMM instrument set**	**Restricted GMM instrument set**
L.ln(LE)	0.5921 (0.000)***	0.6136 (0.000)***	0.6175 (0.000)***	0.6338 (0.000)***
L.ln(YC)	0.0275 (0.002)***	0.0258 (0.022)**	0.0255 (0.010)***	0.0243 (0.041)**
ln(CPC)	0.0526 (0.007)***	0.0498 (0.023)**	0.0507 (0.012)**	0.0488 (0.035)**
L.ln(PEN)	0.0544 (0.001)***	0.0524 (0.006)***	0.0510 (0.001)***	0.0492 (0.012)**
MAL	-0.0294 (0.023)**	-0.0276 (0.062)*	-0.0272 (0.032)**	-0.0259 (0.071)*
HIV	-0.0131 (0.000)***	-.0125 (0.000)***	-0.0129 (0.000)***	-.0124 (0.000)***
Wald	Chi2[12]=11984 (0.000)***	Chi2[12]= 11969 (0.000)***	Chi2(18)=37885 (0.000)***	Chi2(18) = 30424 (0.000)***
Hansen (J-test)	Chi2[26]=38.7 (0.053)*	Chi2[19]=22.32 (0.269)	Chi2(44)=48.95 (0.281)	Chi2(34)=23.32 (0.916)
Difference-in-Sargan (C-test)	Chi2[6]=11.3 (0.081)*	Chi2[5]=4.53 (0.50)	Chi(12)=9.42 (0.667)	Chi2[9]=5.37 (0.801)

Dependent variable: ln(TEN)	**Full sample**	
	Standard GMM instrument set	**Restricted GMM instrument set**
L.ln(TEN)	0.3720 (0.000)***	0.8049 (0.000)***
L.ln(SEN)	0.3310 (0.000)***	-0.0506 (0.599)
L.ln(YC)	0.1428 (0.013)**	0.0806 (0.064)*
ln(LE)	0.6895 (0.021)**	0.3869 (0.081)*
URB	0.0079 (0.002)***	0.0014 (0.464)
HIV	-0.0157 (0.007)***	-.0126 (0.002)***
Hansen (J-test)	Chi2[30]=39.44 (0.116)	Chi2[23]=25.64 (0.318)
Difference-in-Sargan (C-test)	Chi2[6]=3.72 (0.714)	Chi2[5]=6.55 (0.257)

Notes:

1. *Significant at 10%; **Significant at 5%; ***Significant at 1%. Standard errors are robust to arbitrary autocorrelation and heteroskedasticity within countries.
2. L. denotes lagged value.
3. Hansen denotes the Hansen (1982) J-test for over-identifying restrictions, whilst the difference-in-Sargan C-test is a test for the validity of the additional instruments used in system-GMM relative to difference-GMM.
4. Values in parenthesis indicate p-values, that is, the probability of incorrectly rejecting the null hypothesis.
5. The standard instrument set underlying the lagged dependent variable includes lags one and earlier of the instrumenting variable in levels for the equation in differences and lags zero and earlier of the instrumenting variable in differences for the levels equation. The alternative instrument set includes lags two and earlier of the instrumenting variable in levels for the equation in differences and lags one and earlier of the instrumenting variable in differences for the levels equation.

Table 14: Observed coefficients on lagged dependent variables under alternative estimation methods

	OLS	Fixed Effects	Difference-GMM (one-step)	Difference-GMM (two-step)	System-GMM (one-step)
Health – Full sample	0.8925	0.7767	0.4565	0.4993	0.5921
Health – Developing world	0.8969	0.7784	0.4445	0.4935	0.6175
Health – Sub-Saharan Africa	0.8873	0.7252	0.5450	0.5769	0.6703
Secondary Education – Full sample	0.7575	0.4544	0.5711	0.5018	0.6369
Secondary Education – Developing world	0.7615	0.4016	0.3545	0.3320	0.4842
Secondary Education – Sub Saharan Africa	0.7274	0.3507	0.3576	0.3615	0.5368
Tertiary Education – Full sample	0.6772	0.2967	0.3258	0.40512	0.3720
Primary Education – Sub-Saharan Africa	0.7822	0.4827	0.5935	0.6371	0.7266
Growth – Full sample (Version 1)[a]	0.9468	0.7790	0.7729	0.6385	0.8793
Growth – Full sample (Version 2)[b]	0.9495	0.7771	0.7145	0.6608	0.8535
Growth – Developing world	0.8833	0.7441	0.7553	0.7200	0.8133
Growth – Sub Saharan Africa (Version 1)[c]	0.9172	0.7480	0.9253	0.9324	0.9282
Growth – Sub Saharan Africa (Version 2)[d]	0.9163	0.8230	0.9020	0.8867	0.9239

Notes:
In each case, estimation is carried out including a full-set of time dummy variables.
[a] Version 1 refers to the specification including secondary education capital, but not, tertiary education capital.
[b] Version 2 refers to the specification including both secondary and tertiary education capital.
[c] Version 1 refers to the specification including primary education capital, but not, tertiary education capital.
[d] Version 2 refers to the specification including both primary and secondary education capital.

References

Aghion, P. and S. N. Durlauf, Eds. (2005). Handbook of Economic Growth. Volume 1A. Amsterdam and San Diego, Handbooks in Economics vol. 22. Elsevier North-Holland.

Aghion, P. and S. N. Durlauf, Eds. (2005). Handbook of Economic Growth. Volume 1B. Amsterdam and San Diego, Handbooks in Economics vol. 22. Elsevier North-Holland.

Aghion, P. and P. Howit (1998). Endogenous Growth Theory. Cambridge, MA, MIT Press.

Agiomirgianakis, G., D. Asteriou, et al. (2002). "Human Capital and Economic Growth Revisited: A Dynamic Panel Data Study." International Advances in Economic Research 8(3): 177-87.

Ainsworth, M. and M. Over (1994). "AIDS and African Development." World Bank Research Observer 9(2): 203-40.

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

Arndt, C. (2006). "HIV/AIDS, human capital, and economic growth prospects for Mozambique." Journal of Policy Modeling 28: 477-489.

Arndt, C. and J. D. Lewis (2000). "The macro implications of HIV/AIDS in South Africa: A preliminary assessment." South African Journal of Economics 68(5): 856-887.

Baldacci, E., B. J. Clements, et al. (2004). Social Spending, Human Capital, and Growth in Developing Countries: Implications for Achieving the MDGs, International Monetary Fund IMF Working Papers 04/217.

Baltagi, B. H. (2001). Econometric Analysis of Panel Data. New York, John Wiley.

Baltagi, B. H. (2005). Econometric Analysis of Panel Data. Chichester, John Wiley & Sons Ltd.

Barro, R. J. (1996). Health, Human Capital and Economic Growth. Paper for the Program on Public Policy and Health. Washington, Pan American Health Organization and World Health Organisation.

Barro, R. J. and X. Sala-i-Martin (2004). Economic Growth. Cambridge and London, MIT Press.

Bell, C., S. Devarajan, et al. (2006). "The Long-Run Economic Costs of AIDS: A Model with an Application to South Africa." THE WORLD BANK ECONOMIC REVIEW 20(1): 55-89.

Bloom, D. E. and A. S. Mahal (1997). "Does the AIDS Epidemic Threaten Economic Growth?" Journal of Econometrics 77(1): 105-24.

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

Blundell, R. and S. Bond (2000). "GMM Estimation with persistent panel data: an application to production functions." Econometric Reviews 19(3): 321 - 340.

Bollinger, L. and J. Stover. (1999). "The Economic Impact of AIDS." Futures Group International Policy Project, from http://www.iaen.org/impact/stovboll.pdf.

Bond, S., A. E. Hoeffler, et al. (2001). GMM Estimation of Empirical Growth Models. Economics Papers. 2001-W21. Economics Group. Nuffield College, Oxford University.

Bond, S. R. (2002). "Dynamic panel data models: a guide to micro data methods and practice." Portuguese Economic Journal 1(2): 141.

Bonnel, R. (2000). "HIV/AIDS and Economic Growth: A Global Perspective." South African Journal of Economics 68(5): 820-55.
Botswana Institute for Development Policy Analysis (BIDPA) (2000). Macroeconomic Impacts of the HIV/AIDS Epidemic in Botswana. Gaborone, Botswana, BIDPA.
Bowsher, C. G. (2002). "On testing overidentifying restrictions in dynamic panel data models." Economics Letters 77(2): 211-220.
Bureau for Economic Research (BER) (2001). The Macro-Economic Impact of HIV/AIDS in South Africa. Economic Research Note 10, December 2001. Stellenbosch, BER.
Caselli, F., G. Esquivel, et al. (1996). "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics." Journal of Economic Growth 1(3): 363-389.
Corrigan, P., G. Glomm, et al. (2005). "AIDS Crisis and Growth." Journal of Development Economics 77(1): 107-24.
Cuddington, J. T. (1993a). "Modeling the Macroeconomic Effects of AIDS, with an Application to Tanzania." World Bank Economic Review 7(2): 173-89.
Cuddington, J. T. (1993b). "Further Results on the Macroeconomic Effects of AIDS: The Dualistic, Labor-Surplus Economy." World Bank Economic Review 7(3): 403-17.
Cuddington, J. T. and J. D. Hancock (1995). "The Macroeconomic Impact of AIDS in Malawi: A Dualistic, Labour Surplus Economy." Journal of African Economies 4(1): 1-28.
Cuddington, J. T., J. D. Hancock, et al. (1994). "A Dynamic Aggregate Model of the AIDS Epidemic with Possible Policy Interventions." Journal of Policy Modeling 16(5): 473-96.
Culyer, A. J. and J. P. Newhouse, Eds. (2000). Handbook of health economics. Volume 1B. Amsterdam; New York and Oxford, Handbooks in Economics vol. 17.
Elsevier Science North-Holland.
de la Fuente, A. and R. Doménech (2006). "Human Capital In Growth Regressions: How Much Difference Does Data Quality Make?" Journal of the European Economic Association 4(1): 1-36.
Dixon, S., S. McDonald, et al. (2001). "AIDS and Economic Growth in Africa: A Panel Data Analysis." Journal of International Development 13(4): 411-26.
Durlauf, S. N., P. Johnson, et al. (2005). Growth Econometrics. Handbook of Economic Growth. P. Aghion and S. N. Durlauf. Amsterdam, Elsevier B.V. 1A: 555-677.
FAO (1998). FAOSTAT CD-ROM. Food and Agricultural Organisation of the United Nations. Rome.
FAO (2006). FAO Statistical Databases, Food and Agricultural Organisation of the United Nations, Rome. Available: http://faostat.fao.org/site/346/default.aspx.
Fayissa, B. and P. Gutema (2005). "Estimating a health production function for Sub-Saharan Africa (SSA)." Applied Economics 37(2): 155-164.
Ferreira, P. C. and S. Pessoa (2003). The Long-Run Economic Impact of AIDS. Rio de Janeiro, Fundacao Getulio Vargas.
Folland, S., A. C. Goodman, et al. (2004). The Economics of Health and Health Care. New Jersey, Pearson Education International.
Gallup, J., J. Sachs, et al. (1999). Geography and Economic Development, Centre for International Development at Harvard University. Working Paper No. 1. Available: http://www.cid.harvard.edu/ciddata/ciddata.html.
Grossman, M. (1972). "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy 80(2): 223-55.
Gyimah-Brempong, K., O. Paddison, et al. (2006). "Higher Education and Economic Growth in Africa." Journal of Development Studies 42(3): 509-29.

Haacker, M. (2002a). "The Economic Consequences of HIV/AIDS in Southern Africa." IMF Working Paper(W/02/38).
Haacker, M. (2002b). "Modeling the Macroeconomic Impact of HIV/AIDS." IMF Working Paper(W/02/195).
Hall, R. E. and C. Jones (1999). "Why do some countries produce so much more output per worker than others." Quarterly Journal of Economics 114(1): 83-116.
Hansen, L. P. (1982). "Large Sample Properties of Generalized Method of Moments Estimators." Econometrica 50(4): 1029-1054.
Hanushek, E. A. and D. D. Kimko (2000). "Schooling, labor-force quality, and the growth of nations." American Economic Review 90(5): 1184-1208.
Heston, A., R. Summers, et al. (2006). Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. Available: http://pwt.econ.upenn.edu/.
Hill, R. C., W. E. Griffiths, et al. (2001). Undergraduate econometrics. New York :, John Wiley.
Hoeffler, A. E. (2002). "The Augmented Solow Model and the African Growth Debate." Oxford Bulletin of Economics and Statistics 64(2): 135-158.
Islam, N. (1995). "Growth Empirics: A Panel Data Approach." Quarterly Journal of Economics 110(4): 1127-70.
Kambou, G., S. Devarajan, et al. (1992). "The Economic Impact of AIDS in an African Country: Simulations with a Computable General Equilibrium Model of Cameroon." Journal of African Economies 1(1): 103-130.
Keller, K. (2006). "Investment in primary, secondary and higher education and the effects on economic growth." Contemporary Economic Policy 24(1): 18-34.
Knowles, S. and P. D. Owen (1995). "Health Capital and Cross-country Variation in Income Per Capita in the Mankiw-Romer-Weil Model." Economics Letters 48(1): 99-106.
Liu, G. G., J. J. Guo, et al. (2004). "Economic Costs to Business of the HIV/AIDS Epidemic." PharmacoEconomics 22(18): 1181-94.
Lopez, A. D., C. D. Mathers, et al. (2006). "Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data." Lancet 367: 1747-1757.
MacFarlan, M. and S. Sgherri (2001). The Macroeconomic Impact of HIV/AIDS in Botswana. IMF Working Paper. Vol WP/01/80. Washington, IMF.
Mankiw, N. G., D. Romer, et al. (1992). "A Contribution to the Empirics of Economic Growth." Quarterly Journal of Economics 107(2): 407-37.
McDonald, S. and J. Roberts (2002). "Multiple forms of human capital in an augmented Solow model: a panel data investigation." Economics Letters 74(2): 271-276.
McDonald, S. and J. Roberts (2006). "AIDS and economic growth: A human capital approach." Journal of Development Economics 80(1): 228-250.
Mingat, A. and J.-P. Tan (1999). The mechanics of progress in education: Evidence from cross-country data The World Bank Policy Research Working Paper Series (No. 2015), World Bank.
Mingat, A. and J.-P. Tan (2003). "On the mechanics of progress in primary education." Economics of Education Review 22(5): 455-467.
Nelson, R. and E. Phelps (1966). "Investment in Humans, Technological Diffusion, and Economic Growth." American Economic Review 56: 69-75.
Over, M. (1992). The macroeconomic impact of AIDS in sub-Saharan Africa. Population and Human Resources Department. Washington DC, World Bank.

Papageorgiou, C. and P. Stoytcheva (2004). What Do We Know About the Impact of AIDS on Cross-Country Income So Far?, Department of Economics Louisiana State University Departmental Working Papers.
Petrakis, P. E. and D. Stamatakis (2002). "Growth and educational levels: a comparative analysis." Economics of Education Review 21(5): 513-521.
Political Risk Services Group (2006). International Country Risk Guide Online Database. Available: http://www.icrgonline.com/.
Ramcharan, R. (2004). "Higher or basic education? The composition of human capital and economic development." IMF Staff Papers 51(2): 309-326.
Romer, D. (2001). Advanced macroeconomics. Boston, Mass., McGraw-Hill.
Romer, P. (1990). "Endogenous Technological Change." Journal of Political Economy 98(5): S71-S102.
Roodman, D. (2007). "How to do xtabond2: An introduction to "Difference" and "System" GMM in Stata." Center for Global Development Working Papers (Working Paper 103).
Schultz, T. P. (1999). "Health and Schooling Investments in Africa." Journal of Economic Perspectives 13(3): 67-88.
Sianesi, B. and J. Van Reenen (2003). "The Returns to Education: Macroeconomics." Journal of Economic Surverys 17(2): 157-200.
Solow, R. M. (1956). "A contribution to the theory of economic growth." Quarterly Journal of Economics 70: 65-94.
Tandon, A. (2005). "Macroeconomic Impact of HIV/AIDS in the Asian and Pacific Region." Asian Development Bank - ERD Working Paper Series(75).
Temple, J. R. W. (1999). "The new growth evidence." Journal of Economic Literature 37(1): 112-156.
UNAIDS (2006). AIDS Epidemic Update: December 2006. Geneva, UNAIDS.
UNAIDS (2006). Report on the global AIDS epidemic. UNAIDS. Geneva.
UNESCO (2006). UNESCO Education Statistics Online Database, Statistical Branch of the United Nations Educational, Scientific and Cultural Organisation. Available: http://www.uis.unesco.org/en/stats/centre.htm.
Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data. Cambridge, Mass. :, MIT Press.
World Bank (1998). The Initiative on Defining, Monitoring and Measuring Social Capital: Overview and Program Description. Social Capital Initiative Working Paper no. 1. Washington, World Bank.
World Bank (2006). World Development Indicators *Online*. Available: http://www.worldbank.org/data.
World Health Organisation (2001). Macroeconomics and health: investing in health for economic development. Report of the commission on macroeconomics and health. Geneva, WHO.
Wossmann, L. (2003). "Specifying Human Capital." Journal of Economic Surverys 17(3): 239-270.
Young, A. (2005). "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations." Quarterly Journal of Economics 120(2): 423-66.