What do we know about the impact of AIDS on income so far?*

Petia Stoytcheva
Department of Economics
Louisiana State University
Baton Rouge, LA 70803
E–mail: pstoyt1@lsu.edu

May, 2004

Abstract

This paper sheds new light on the impact of AIDS on cross-country income levels. Our empirical analysis uses data for 90 countries spanning the period 1979 to 2000 during which AIDS has spread across the world. We control for a variety of factors that are potentially related to income as suggested by our empirical model and by previous work on economic growth. Using the extended Solow model as our baseline empirical specification, we consider both cross-sectional and panel estimation. It is shown that the number of AIDS cases has a negative and significant effect on the level of income for the full sample in both the cross-sectional and panel estimations. When we arbitrarily split our full sample into OECD and non-OECD countries, we find that the AIDS coefficient continues to be negative and significant for the non-OECD sub-sample, but not for the OECD sub-sample. When we use Hansen's (2000) endogenous splitting methodology, we find that AIDS is a threshold variable that can split countries in our sample into four different regimes, obeying different statistical models. Our results show that each additional AIDS case per $100,000$ persons per year is associated with 0.4 percentage point reduction in the level of output. In addition, we show that the AIDS coefficient on age group 16-34 is negative and significant.

*I thank Professors Chris Papageorgiou, Marios Zachariadis, Andreas Savvides, Carter Hill, Douglas McMillin, Benjamin Bridgman, Douglas Schwalm, Tibor Besedes, participants at the Southern Economic Association 2003 meetings and LSU seminar series for their comments and suggestions. I am also grateful to Txema Calleja from the World Health Organization for providing me with the AIDS definition from "Weekly Epidemiological Record", Geneva (1994) and other AIDS related materials.

1 Introduction

The World Heath Organization (WHO) estimated that in December 2002, 42 million people were living with the human immunodeficiency virus (HIV) or the acquired immune deficiency syndrome (AIDS). The newly infected with HIV in 2002 totaled 5 million and AIDS related deaths in 2002 were 3.1 million. HIV/AIDS now ranks as the world's fourth largest cause of death, after heart disease, strokes and acute lower respiratory infections (Dixon, McDonald, Roberts (2002)). The most affected continent is Africa and particularly sub-Saharan Africa. Although AIDS is a relatively recent infectious disease (spreading across the world over the last two decades,) it is feared that it will soon surpass malaria, which has been around for at least a millennium, as the most deadly infectious disease.

AIDS' alarming infection rate coupled with no known cure may have very important social, political, demographic and economic implications. A central point of analysis for economists is to evaluate the aggregate impact of AIDS on economic growth and welfare. However, a definitive answer to the question regarding the qualitative impact of AIDS on income has yet to be given. Nor has the aggregate quantitative impact of AIDS been empirically evaluated to this date.

Several theoretical papers have addressed this question and have suggested possible negative consequences of the pandemic. For example, Cuddington (1993), simulating a modified Solow model environment, suggests that AIDS, via its impact on morbidity and mortality rates, might reduce GDP in Tanzania in 2010 by 15 to 20 percent relative to a counterfactual no-AIDS scenario. Similarly, Cuddington and Hancock (1994) simulated the impact of AIDS on the Malawian economy and suggest that the average annual real per capita GDP growth over the 1985-2010 period might be $0.2-0.3$ percentage points lower compared to the no-AIDS case. More recently, Ferreira and Pessoa (2003) propose a model in which AIDS impacts negatively on income by affecting the incentives to study due to shorter expected longevity. Based on their model, the most affected countries in sub-Saharan Africa are predicted to become about 25 percent poorer than they would be without AIDS, with schooling declining by about 50 percent. Finally, Corrigan, Glomm, and Mendez (2003) propose a model consistent with large growth effects of the AIDS epidemic, mainly through the effect of the drop in life expectancy on investment and the large generation of orphans created by AIDS.

Even though the above papers have contributed to our understanding of the problem, they are based on theoretical models and numerical simulation exercises and do not utilize actual data on AIDS incidences. In an important contribution, Bloom and Mahal (1997) find that "... there is more flash than substance to the claim that AIDS impedes national economic growth." The paper's main argument is that the AIDS epidemic has had an insignificant effect on the growth rate of per capita income, with no evidence of reverse causality. The authors use standard epidemiological models to estimate the numbers of AIDS cases from information on HIV prevalence at a point in time. Thus, these papers fail to empirically evaluate the aggregate qualitative and quantitative impact of AIDS on growth. With the passage of time and extensive efforts by international organizations like WHO and UNAIDS, however, emerging data on HIV/ AIDS has made empirical investigation of the impact of AIDS more feasible. Using cross-country regressions for the 1990-1997 period, Bonnel (2000) shows that HIV/AIDS has reduced the rate of growth of Africa's per capita income by 0.7 percentage points per year. For those countries affected by malaria, growth was further lowered by 0.3 percentage points per year.

This paper aims to extend the empirical work on the impact of AIDS by asking the following question: what do we know about the impact of AIDS on income so far? Although AIDS is a recent infectious disease, spreading across the world in the last two decades, its negative impact has already been felt in sub-Saharan Africa. AIDS is erasing decades of progress made in extending life expectancy. Average life expectancy in sub-Saharan countries is now 47 years , when it could have been 62 without AIDS. In order to address the economic implications of the disease on welfare, our framework focuses on levels rather than growth of output. A recently proposed alternative to growth regressions has been "levels" regressions (Hauk and Wacziarg, 2004). Levels capture the differences in long-run economic performance (Hall and Jones,1999). A second argument is that differences in growth rates across countries may be mostly transitory.[1] By focusing on level regressions, we try to explain the variation in income across countries and, more importantly, the effect of AIDS on economic performance. In a recent paper Easterly and Levine (2000) document five stylized facts on economic growth and argue that they are more consistent with technology explanation of growth and income differences rather than a total factor accumulation.[2] After accounting for physical and human capital accumulation, there is "something else", that accounts for the difference in income across countries - Total Factor Productivity (TFP). There are different theories, providing different

[1]Other examples that use level regressions are Frankel and Romer (1999) and Acemoglu, Johnson and Robinson (2001).

[2]For more details, see Easterly and Levine (2000).

views about TFP (for example, some of them model TFP as changes in technology, other highlight the role of externalities). In this context, AIDS can be viewed as a productivity shock, affecting individuals in their most productive years of their lives.

First, we make use of the officially reported AIDS cases by the WHO for the period 1979-2000 across 117 countries.[3] This enables us to consider both cross-sectional regression and panel techniques to study the impact of the disease on the level of income. Second, making use of PWT 6.1 we extend the Mankiw, Romer and Weil (1992) (MRW hereafter) dataset until the year 2000 and obtain results for the augmented (with human capital) Solow model for the period 1979-2000. Then we merge the extended MRW dataset with the AIDS dataset and obtain results using regressions based on the extended Solow model with AIDS as an additional explanatory variable.[4] Third, we employ the data splitting methodology proposed by Hansen (2000) to examine whether AIDS is a valid threshold variable that can cluster countries into groups with distinct growth paths.

Our main findings are as follows: First, we show that the number of AIDS cases has a negative and significant effect on the level of income for the full sample in both the cross-sectional and panel estimations. When we arbitrary split our full sample into OECD and non-OECD countries, we find that the AIDS coefficient continues to be negative and significant for the non-OECD sub-sample, but not for the OECD sub-sample. Second, when we use Hansen's (2000) endogenous splitting methodology, we find that AIDS is a threshold variable that can split countries in our sample into four different regimes, obeying different statistical models. Our results show that each additional AIDS case per $100,000$ persons per year is associated with 0.4 percentage point reduction in the level of output. Third, we show that the AIDS coefficient on age group 16-34 is negative and significant.

The remainder of the paper is organized as follows. Section 2 briefly presents the extended Solow model from which we obtain our regression equation. Section 3 describes the data used in our analysis and discusses several aspects of the AIDS dataset in detail. Section 4 discusses our cross-sectional and panel estimation results using the full sample, the arbitrarily and endogenously chosen sub-samples and the four age groups. Section 5 concludes.

[3]In our empirical exercise we use the non-oil sample of Mankiw, Romer and Weil, which consists of 98 countries, but our sample was reduced to 90 countries due to availability of AIDS data.

[4]Following Gallup and Sachs (2000) and McCarthy, Wolf and Wu (2002), we include AIDS in our regression exercise, not log(AIDS).

2 The Extended Solow Model

We start by reviewing the basic and extended Solow model following MRW. In the basic Solow model we assume a Cobb-Douglas production function with two inputs, capital and labor, given by

$$Y_i = K_i^{\alpha}(AL_i)^{1-\alpha},$$

where Y_i is the aggregate output in country i, K_i is the aggregate capital stock, AL_i technology-augmented labor and α is the share of capital. A and L grow exogenously at rates g and n, respectively, and $\alpha \in (0,1)$.

The importance of human capital to economic growth was noticed by many economists, and human capital has been incorporated in many theoretical models and empirical analyses. MRW point out that at the empirical level, the existence of human capital can alter the analysis of cross-country differences. With human capital (H_i) as a third input, the Cobb-Douglas production function is given by

$$Y_i = K_i^{\alpha} H_i^{\beta} \left(AL_i\right)^{1-\alpha-\beta},$$

where $\beta \in (0,1)$ is the share of human capital. Physical and human accumulation equations take the forms $dK_i/dt = s_{ik}Y_i - \delta K_i$ and $dH_i/dt = s_{ih}Y_i - \delta H_i$, respectively, where δ is assumed to be an identical depreciation rate for capital, and s_{ik} and s_{ih} are the fractions of income invested in physical and human capital, respectively. After solving for the steady-state, we obtain *the extended Solow equation*

$$\ln\left(\frac{Y_i}{L_i}\right) = \ln A\left(0\right) + gt + \frac{\alpha}{1-\alpha-\beta}\ln\left(\frac{s_{ik}}{n_i+g+\delta}\right) + \frac{\beta}{1-\alpha-\beta}\ln\left(\frac{s_{ih}}{n_i+g+\delta}\right). \quad (1)$$

3 A First Look at the Data

We begin by describing the data used in our estimation. First, we explain how we update the MRW original dataset. More importantly, we explain in detail our AIDS dataset.

3.1 Extending the MRW dataset

We extend the MRW original dataset (PWT version 4.0) until the year 2000, using PWT version 6.1 for the non-oil sample consisting of 98 countries. Due to the data constraints with the AIDS dataset, our sample was reduced to 90 countries. Our dependent variable is the log (GDP per working age population) in 2000. Data on the percentage of the working age population (aged 15 to 64) is from the World Bank's *World Development Indicators* (2002).[5] We average the population growth of the working-age population n for the period 1979-2000 and add $g + \delta$, which is assumed to be 0.05. Following MRW, the saving rate s_k is the ratio of average investment to GDP over the 1979-2000 period (PWT 6.1).[6] We add a variable called *School* to proxy for s_h that measures the percentage of the working-age population that is in secondary school. Our human capital measure is taken from Bernanke and Gürkaynak (2001).[7]

3.2 The AIDS dataset

In our estimation we include a new variable - AIDS. We were able to assemble data for 117 countries over 1979-2000, the period during which AIDS as an epidemic has spread across the world. This enables us to examine the impact of the disease not only using cross-country regression, but also using panel-data estimation techniques. We constructed AIDS using the officially reported cases from the WHO/UNAIDS Global Surveillance fact sheets.[8] For each country in the sample we start from the year during which a case was reported. We multiply the number of reported cases times 100,000 and divide by total population in each year (data on population is from the *World Development Indicators* (2002)) to obtain incidence per 100,000 per country per year. The officially reported AIDS cases represent the number of new AIDS infections, occurring each year. Thus, we obtain AIDS incidence, which is a flow measure. For the cross-sectional estimation, for each country in the sample we average the cases, starting from the year in which a case was reported up to the year 2000.[9] For the panel estimation, we average the data into 5 year periods since the disturbance terms are less likely to be influenced by business cycle fluctuations over five year intervals. Thus, we have now three non-overlapping five-year time intervals - from 1985 to 1990, 1990 to 1995, and

[5]The period between $1979 - 2000$ will be the period in consideration in our estimation, because AIDS has spread since 1979.

[6]For more details on the variables used, see the appendix.

[7]Bernanke and Gürkaynak (2001) obtain their human capital measure by multiplying the fraction of the eligible population (age 12-17) enrolled in secondary school times the fraction of the working-age population that is of school (age 15-19). This measure is clearly imperfect, as MRW point because the age ranges in the two data series are not exactly the same but it is not likely to create major biases. We average human capital for the period 1970-1995.

[8]The WHO case definition for AIDS surveillance is from "Weekly Epidemiological Record", WHO, Geneva (1994). For detailed description of the definition, see the appendix.

[9]For details about the sample of countries and relevant variables used in the estimation, see the appendix, Table A1.

Table 1: Descriptive statistics

Continent		Mean	Standard deviation	Minimum	Maximum
Africa	GDP per capita	2,303	2,466	461	10,294
	AIDS	21.839	37.299	0.021	173.043
Americas	GDP per capita	6,192	5,234	1,075	22,934
	AIDS	6.326	6.734	0.217	26.818
Asia	GDP per capita	7,951	6,799	1,004	21,205
	AIDS	1.129	3.596	0.001	17.047
Europe	GDP per capita	15,322	5,595	4,424	29,274
	AIDS	2.046	2.127	0.022	8.412
Oceania	GDP per capita	10,566	7,855	3,152	19,424
	AIDS	1.433	1.120	0.162	2.872
World	**GDP per capita**	**7,149**	**6,858**	**461**	**2,9274**
	AIDS	**9.800**	**24.169**	**0.001**	**173.043**

Notes: The mean, the standard deviation, the minimum and the maximum values presented above are computed for 42 countries in Africa, 25 countries in the Americas, 22 countries in Asia, 25 countries in Europe, 4 countries in Oceania. Growth GDP and AIDS are averages since an AIDS case was reported annually from 1979 till 2000.

1995 to 2000.

Next we present descriptive statistics of the AIDS data which reveal important new information.[10] Table 1 presents the mean, standard deviation, minimum and maximum of AIDS and GDP per capita for all the continents. We average GDP per capita for the period starting from the year in which an AIDS case was reported till 2000 (PWT 6.1). We note that the mean for AIDS in Africa (21.839) is much higher than in all other continents. These are countries heavily affected from the epidemic, particularly sub-Saharan Africa.[11] Another interesting thing to notice is the very high incidence of AIDS in the Americas (the mean is 6.326). It is much higher than in Europe, where the mean incidence of AIDS is 2.016. Possible explanations for the high rates of infection in those countries in the Americas are religion, income, education. We present the correlations of these two variables for Africa and Europe in the appendix (Figures 5 and 6). We find that the correlation between AIDS and GDP per capita in Africa is -0.017, while the same correlation in Europe is 0.220.

[10]We group countries into continents to examine whether geographical location matters.

[11]Our sample includes the following sub-Saharan countries: Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Chad, Congo, Ghana, Kenya, Malawi, Mali, Mauritius, Mozambique, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Tanzania, Togo, Uganda, Zambia, Zimbabwe.

4 Estimation and Results

In this section we first present the cross-sectional results for the full sample and arbitrarily chosen sub-samples as well as endogenously chosen sub-samples. Then we perform robustness analysis to potential outliers and we examine the effect of AIDS on the different age groups, for which we were able to assemble data. Next, we extend the analysis to panel estimation and discuss implications of the results to the existing literature on AIDS and income.

4.1 Cross-sectional estimation

Table 2 presents estimates for the *extended Solow model* for the period 1979 till 2000 for the full sample and arbitrarily chosen OECD and non-OECD subsamples using ordinary least squares (OLS). First, we estimate the MRW specification with our extended data and then we add AIDS as a regressor. The dependent variable in the regression equation is the ln(GDP per Worker in 2000).[12]

When we replicate the *extended Solow* using PWT 6.1 for the full sample of 90 countries, we find that the model (column 2) explains 76% of the overall variation in per capita income. Adding AIDS into the regression, the variables can explain 77% of the overall variation in per capita income (column 5). The estimates from the models have the expected signs, but there is a difference in terms of the magnitudes. The estimated coefficient for physical capital decreases from 0.8544 to 0.8312, keeping the same significance level at 1%. Similarly, the coefficient on human capital remains significant in both models at 1%. The estimated coefficient on $\ln(n_i + g + \delta)$ is -0.5756 in the *extended Solow* and increases to -0.5435 in the *extended Solow* with AIDS, remaining highly significant at 5%. These results are consistent with MRW when they estimate the model using data for the period 1960-1985. They are also similar to Bernanke and Gürkaynak (2001), who extend the data till 1995, using PWT 6.0. The estimated coefficient on AIDS is negative and significantly different from zero at 5%. This shows that AIDS has a negative impact on output. Each additional AIDS case per $100,000$ persons per year is associated with a 0.4 percentage point reduction in output.

Next, we examine the robustness of our results by arbitrarily splitting the full sample into OECD and non-OECD countries. For the non-OECD countries, we obtain a positive and significant

[12] We repeated the estimation using utput per capita in 2000 as a dependent variable and the results were qualitatively similar to those, obtained with output per worker population in 2000.

Table 2: Cross-country regressions for the full sample and for OECD and non-OECD countries

Dependent variable: ln(GDP per Worker in 2000)						
Specification	Extended Solow Model (*PWT 6.1*)			Extended Solow Model with AIDS (*PWT6.1*)		
	Non-oil	OECD	Non-OECD	Non-oil	OECD	Non-OECD
Unrestricted						
Constant	11.2520 ***	15.8540***	11.4649 ***	11.3217***	15.3411***	11.5315***
	(1.0290)	(1.7752)	(0.8743)	(0.9958)	(1.6330)	(0.8135)
$\ln s_{ik}$	0.8544***	1.1201**	0.6104***	0.8312***	1.0287**	0.5904***
	(0.1391)	(0.4782)	(0.1481)	(0.1364)	(0.4150)	(0.1474)
$\ln(n_i + g + \delta)$	−0.5756*	0.3892 *	−0.1592	−0.5435*	0.2752	−0.1313
	(0.3371)	(0.1951)	(0.3417)	(0.3249)	(0.2761)	(0.3180)
$\ln s_{ih}$	0.6820***	1.1929**	0.6219***	0.6807***	1.1876 **	0.6214***
	(0.1045)	(0.4782)	(0.0976)	(0.1000)	(0.4448)	(0.0941)
AIDS				−0.0043**	0.0162	−0.0039**
				(0.0018)	(0.0198)	(0.0020)
Adj. R^2	0.76	0.45	0.67	0.77	0.48	0.67
Obs.	90	21	69	90	21	69

Notes: Standard errors are in parentheses. It is assumed that $g+\delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level. The investment and population growth rates are averages for the period 1979-2000. s_h is the average percentage of the working-age population in secondary school for the period 1970-1995.

coefficient on $\ln(s_{ik})$-0.6104, a positive and highly significant from coefficient on $\ln(s_{ih})$-0.6219, and negative and insignificant coefficient on $\ln(n_i + g + \delta) - 0.1592$. After we include AIDS in the regression, the model still explains 67 % of the overall variation in per capita income and there is little change in the results (column 7) compared with the results from estimating the non-OECD subsample without AIDS as a regressor. The non-OECD sample consists mainly of countries, where most people infected cannot afford treatment and since AIDS affects people in these countries in their most productive years, it is expected to adversely affect the level of income per capita.

When we compare the *extended Solow model* for the OECD (column 3) with MRW estimates of the model for the period 1960-1985, we find a significant improvement in terms of the fit of the model. Our data can explain 45% of the overall variation in per capita income for the OECD countries, which is relatively high compared to 24% in MRW. There is also a difference in terms of the estimated coefficients. In our estimation we obtain higher estimated coefficients for physical capital, human capital and for $(n_i + g + \delta)$. After adding AIDS in the regression, the coefficient on s_{ik} decreases from 1.1201 to 1.0287, remaining the same significance level at 5%. The coefficient on s_{ih} remains almost the same in terms of magnitude and significance level (column 6). The estimated coefficient on $n_i + g + \delta$ is significant at 5% in the model without AIDS, but becomes insignificant when we include AIDS as a regressor (column 6). In the OECD sample, the estimated coefficient on AIDS is insignificant, showing that the epidemic has no significant impact on the level of income for these countries.

A possible explanation for this result is the introduction of the antiretroviral drugs. Since people in rich countries can afford treatment, it can delay the transmission of the disease and may cause positive externalities by protecting other people. However, the impact of antiretrovirals on the spread of the epidemic is unclear (Kremer (2002)). Advocates of antiretroviral drugs for HIV/AIDS support the view that the effect of the drugs is expected to be positive, leading to enlarge prevention and slow transmission. But even if the availability of the drugs encourages testing, once people are infected from the virus, they could choose to have more sexual contacts knowing they could lose nothing. In the developing countries, the effect of the pandemic is different. People cannot afford the expensive drugs and because of the very low level of education, they are not even familiar with the basic protection measure-the use of a condom. Sebnem Kalemli-Ozcan (2001) provides new evidence on the empirical relationship between the mortality changes and the quality-quantity trade-off for a panel of African countries, where parents will choose to have more children and provide them with less education facing a high probability of getting infected with the epidemic. In a recent paper Papageorgiou, Savvides and Zachariadis (2004) show that medical technology developed in countries close to the technology frontier have a significant impact on health and income in countries distant from this frontier. This is true as well for antiretroviral drugs for HIV/AIDS, which, imported from developed countries into developing countries may have a substantial positive effect by lowering the infection rates.

In addition to the level regressions, we examine the effect of the disease on the growth of real GDP per worker for the period 1979-2000. We present the results in the appendix - Table 8.

When we estimate the model for the full sample without AIDS, we find that our model explains 36% of the total variation in growth of GDP per worker. The estimate on initial income enters negative and insignificant, the estimate on $\ln(s_{ik})$ is positive and significant at the 1% level, the coefficient on $\ln(n_i + g + \delta)$ is negative and insignificant, and the estimate on human capital is positive and highly significant. When we arbitrarily split the sample into OECD and non-OECD countries, we notice a change in terms of the significance level and magnitude of the coefficients. For the OECD sample, the estimate on $\ln(s_{ik})$ is insignificant, the estimate on $\ln(n_i + g + \delta)$ is positive and significant at the 5% level and the coefficient on human capital is positive and significant. The estimates on investment, population growth and human capital are similar to the estimates obtained for the full sample (column 2). After we include AIDS in the regression, the AIDS coefficient is negative and insignificant for the full sample and non-OECD countries and positive and significant for the OECD sample. This shows that AIDS has an insignificant impact on the growth rate of GDP per worker. In the non-oil sample, the estimate on initial output is negative and insignificant, the estimate on $\ln(s_{ik})$ is positive and significant at the 10% level, the estimate on $\ln(n_i + g + \delta)$ becomes more negative, compared to the model without AIDS. The coefficient on $\ln(s_{ih})$ increases in terms of the magnitude, compared to the model without AIDS and changes its significance level to 5%. The estimates for the non-OECD sub-sample are similar to the model without AIDS (column 7). However, for the OECD sub-sample we notice a significant change in terms of the magnitude and significance level for the estimates on $\ln(s_{ik})$ and $\ln(n_i + g + \delta)$ compared to the model without AIDS. The coefficient on $\ln(s_{ik})$ becomes negative, insignificant and the coefficient on population growth losses its significance level, but remains positive.

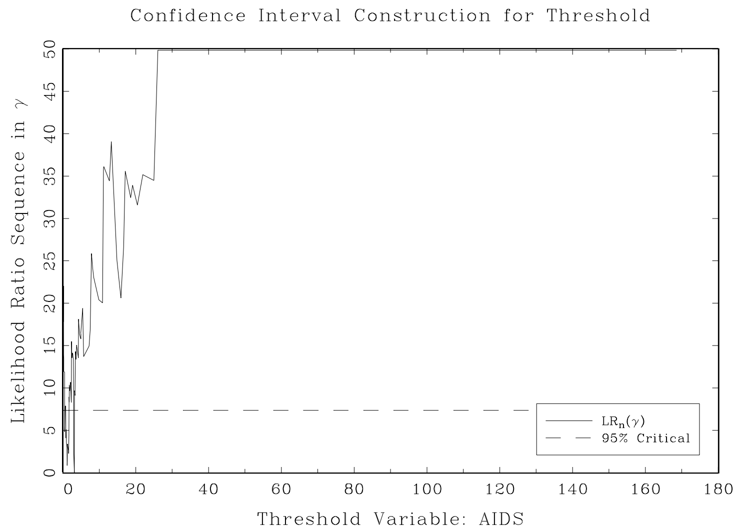
4.2 Threshold estimation

In this section we follow Hansen (2000) and allow the data to endogenously select regimes using different variables, rather than arbitrarily splitting the sample as in the previous section. The advantage of Hansen's methodology over the regression-tree methodology used in Durlauf and Johnson (1995) is that it is based on an asymptotic distribution theory. We allow the data to endogenously select regimes using AIDS. We tried to split our data using initial output, initial schooling and initial AIDS. The bootstrap p-values were significant in the first round of splitting, showing that there might be a split[13]. We chose average AIDS[14] as a possible threshold variable, since AIDS is a dynamic variable and captures the changes in the progression of the disease for each country in our sample. Since the AIDS epidemic has started since 1979, we use the period

[13] The programs are available upon request.

[14] AIDS is the average incidence per 100,000 per country for the period 1979-2000.

Figure 1: First Sample Split



1979-2000 and the data from our cross-sectional estimation for the same period.

We consider the following regression equation:

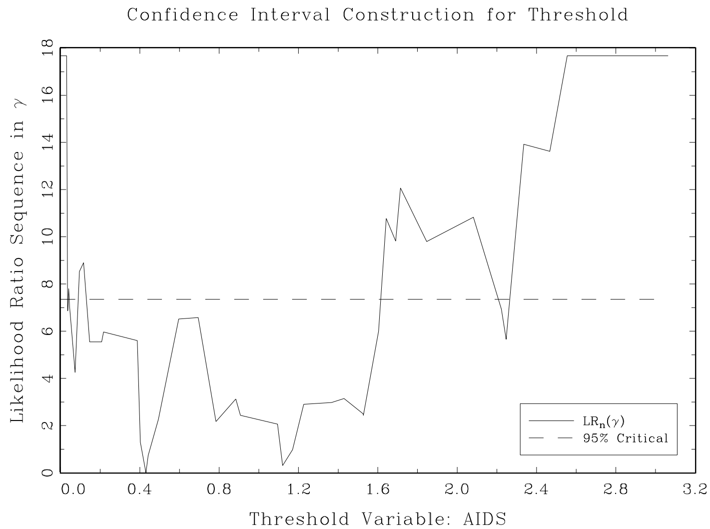
$$\ln y_{i,2000} = a_0 + a_1 \ln s_{i,k} + a_2 \ln(n + g + \delta) + a_3 \ln s_{i,h} + a_4 \text{AIDS}_i + \varepsilon_i, \tag{2}$$

where y_i is output per working age person in country i, s_{ik} is the ratio of average investment to GDP over 1979-2000, s_{ih} is secondary school enrollment of working-age population, n is average population growth, $g + \delta = 0.05$ as in MRW, AIDS_i is the incidence per 100,000 per country, averaged for the period 1979-2000.

In the first round of splitting, we find the following p-value: 0.001 for AIDS, which shows that there may be a sample split based on AIDS. Figure 1 presents the normalized likelihood ratio sequence $LR_n^*(\gamma)$ statistic as a function of AIDS threshold. The least-squares estimate γ is the value that minimizes the function $LR_n^*(\gamma)$ which occurs at $\gamma = 3.0637$. The asymptotic 95% critical value (7.35) is shown by the dotted line and where it crosses $LR_n^*(\gamma)$ displays the confidence set $[0.43136, 3.0637]$. AIDS as a threshold divides the full sample (90 countries) into two sub-samples: one, containing 51 countries (AIDS < 3.0637) and second, which comprises 39 countries with higher incidence of AIDS (AIDS > 3.0637).[15]

[15] We present the regression-tree diagram with the two regimes, the countries from each regime and the cross-sectional estimation for the two regimes in the appendix: Figure 7, Tables 9 and 10, correspondingly.

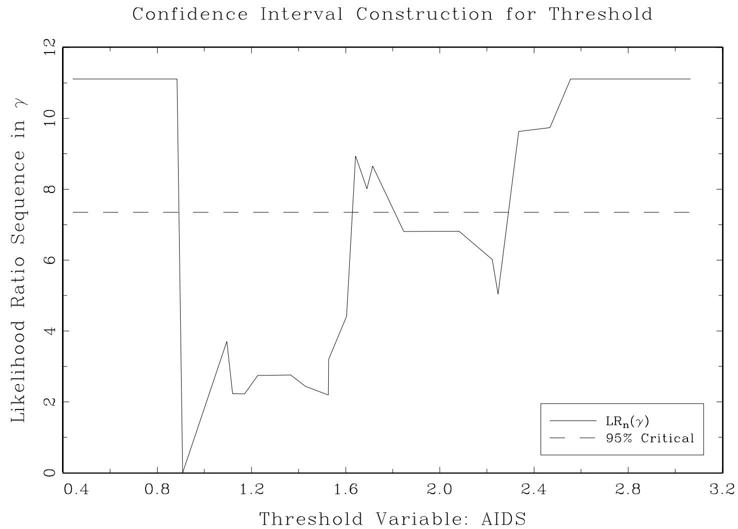
Figure 2: Second Sample Split



We tried to further split the group with the higher AIDS incidence, but the bootstrap test statistic was insignificant. However, the bootstrap test statistic for the sample with 51 countries was significant (0.026), showing a possible sample split. Figure 2 presents the normalized likelihood ratio sequence $LR_n^*(\gamma)$ statistic as a function of AIDS threshold. The least-squares estimate γ is the value that minimizes the function $LR_n^*(\gamma)$ which occurs at $\gamma = 0.4314$. The asymptotic 95% critical value (7.35) is shown by the dotted line and where it crosses $LR_n^*(\gamma)$ displays the confidence set $[0.03599, 2.24708]$. AIDS as a threshold divides our sub-sample of 51 countries into two groups: one, which consists of 19 countries with AIDS incidence lower than 0.4314 and another, that contains 32 countries with incidence higher than 0.4314.

We tried to further split the group with AIDS incidence lower than 0.4314, but the bootstrap test statistic was insignificant. The bootstrap test statistic for the group with incidence higher than 0.4314 is significant at 10%, which shows that there is a possible split. Figure 3 presents the normalized likelihood ratio sequence $LR_n^*(\gamma)$ statistic as a function of AIDS threshold. The least-squares estimate γ is the value that minimizes the function $LR_n^*(\gamma)$ which occurs at $\gamma = 0.9070$. The asymptotic 95% critical value (7.35) is shown by the dotted line and where it crosses $LR_n^*(\gamma)$ displays the confidence set $[0.90704, 2.24708]$. AIDS as a threshold divides our sub-sample of 32 countries into two groups: one, which consists of 7 countries with AIDS incidence lower than 0.9070

Figure 3: Third Sample Split



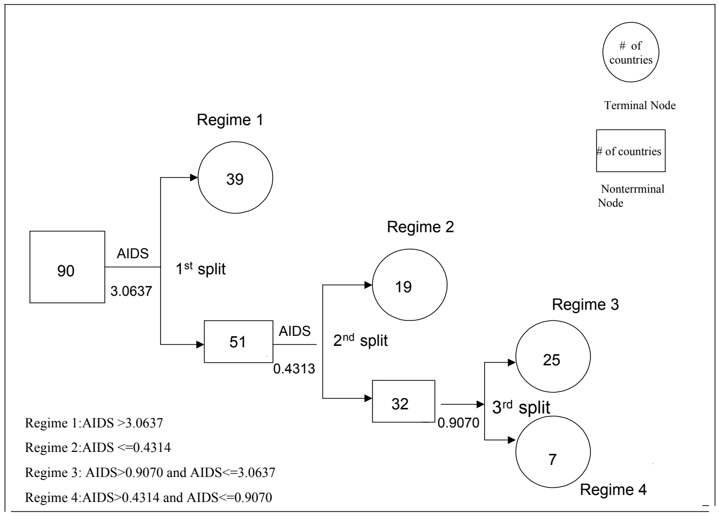
and another, that contains 25 countries with incidence higher than 0.9070.

Figure 4 presents the regression tree diagram. Non-terminal nodes are illustrated by squares where terminal nodes are illustrated by circles. The numbers insidethe squares and circles show the number of countries in each node. The point estimates for the threshold variable are presented on the rays connecting nodes. Table 3 shows the countries from the four regimes.

Table 4 presents the estimates from the cross-country regression of the four regimes. Regime 1 consists of 39 countries with incidence of AIDS higher than 3.0637. Most of the countries are sub-Saharan countries, highly infected from the epidemic. These are at the same time poor countries, for which the epidemic is expected to have a negative impact on income. The coefficient on AIDS is -0.0042, significant at the 10% level. The coefficients on $\ln(\mathrm{s}_{ik})$ and $\ln(n_i + g + \delta)$ are positive, significant at the 5% and 10% respectively. The estimate on $\ln(n_i + g + \delta)$ enters insignificant.

Regime 2 consists of 19 countries with incidence of AIDS lower than 0.4314. The estimate on AIDS is insignificant, showing that AIDS has no quantifiable impact on the income level. The estimates on $\ln(s_{ik})$, $\ln(n_i + g + \delta)$, and $\ln(s_{ih})$ are significant at the 1% level (column 3). When we estimate the model for Regimes 3 and 4, we notice a big variation in terms of the magnitude and significance level of the estimates. The coefficient on AIDS enters insignificant in both regimes (columns 5 and 6). The estimate on human capital is positive and significant at the 1% level

Figure 4: Regression Tree Diagram



for Regime 3, but insignificant for Regime 4. The estimate $\ln(n_i + g + \delta)$ is insignificant in the regression for the two regimes, and the estimate on ln s_{ih} is positive and significant. We can interpret these results as an evidence of parameter heterogeneity, which shows that countries can be grouped according to different statistical models. Recent papers by Durlauf and Johnson (1995), Papageorgiou (2002), Masanjala and Papageorgiou (2004) provide strong evidence of the existence of multiple regimes. More importantly, we show that AIDS is a threshold variable, that can group countries into different regimes. This shows that not only income plays a crucial role for AIDS, but other factors like religion, culture, institutions, geography are important. Acemoglu, Johnson and Robinson (2002) argue that the main impact of disease environments on the economic development of nations is due to indirect effect of health conditions on income through institutions. This is true for AIDS as well, although as a disease it does not have a long history, the evolution of AIDS is determined to large extend from factors that affect the development paths of the countries. One of them is institutions.

4.3 Robustness analysis

In this section we examine the robustness of our results to excluding outliers from the sample. Our sample has five severe outliers - Botswana, Congo, Malawi, Zimbabwe, Zambia.[16] These are sub-

[16]These outliers are exluded only when we consider the variable of interest, AIDS.

Table 3: Countries from the four regimes

Regime1		Regime 2	Regime 3		Regime 4
Angola	Kenya	Algeria	Argentina	S. Africa	Ecuador
Benin	Malawi	Bangladesh	Australia	Sweden	Hong Kong
Botswana	Mali	Bolivia	Austria	U. K.	Israel
Brazil	Mozambique	Egypt	Belgium	Uruguay	Norway
Burkina Faso	Niger	Finland	Chile	Venezuela	Paraguya
Burundi	Nigeria	India	Canada		Sierra Leone
Cameroon	Panama	Indonesia	Columbia		Tunisia
C. Africa	Portugal	Japan	Denmark		
Chad	Rwanda	Jordan	Greece		
Congo	Spain	Korea	Guatemala		
Costa Rica	Switzerland	Madagascar	Ireland		
Dom. Rep.	Tanzania	Mauritius	Malaysia		
El Salvador	Thailand	Morocco	Mauritania		
Ethiopia	Togo	Nicaragua	Mexico		
France	Tr.&Tobago	Pakistan	Netherlands		
Ghana	Uganda	Philippines	N.Zealand		
Haiti	USA	Sri Lanka	Papua N. G.		
Honduras	Zambia	Syrian Arab Rep.	Peru		
Italy	Zimbabwe	Turkey	Senegal		
Jamaica			Singapore		
(39)		(19)	(25)		(7)

Saharan countries with the highest concentration of the epidemic. This again raises the question of what factors have contributed to the extremely high incidence of AIDS specifically in that region. In a recent paper Masanjala and Papageorgiou (2004) examine whether the determinants of economic growth for Africa are different from the rest of the world, using Bayesian Model Averaging (BMA) methodology. The paper's main argument is that the determinants of economic growth in Africa are strikingly different from the rest of the world. Not surprisingly, the outliers in our sample are from sub-Saharan Africa. We reestimated the augmented Solow equation without the outliers for both the full sample and the non-OECD sample. We present the results in Table 11 in the appendix.

For the full sample, which consists of 85 countries after excluding the potential outliers, we obtain negative and insignificant estimate on AIDS. A possible explanation for this result is the fact that these five countries have the highest incidence of AIDS. Since our full sample includes

Table 4: Cross-country regressions for the four regimes

Dependent variable: ln(GDP per worker in 2000)				
Specification	Regime 1	Regime 2	Regime 3	Regime 4
Constant	9.3328***	5.9007***	12.4614***	9.6545***
	(2.7346)	(1.6260	(0.7695)	(2.1073)
ln s_{ik}	0.5474**	0.4714***	0.7116***	2.9878
	(0.2023)	(0.1477)	(0.2100)	(1.1456)
$\ln(n_i + g + \delta)$	−1.1220	-2.3053***	−0.2380	−0.9589
	(0.8748)	(0.6439)	(0.1860)	(0.7009)
ln s_{ih}	0.7290*	0.8142***	0.8017***	−1.2408
	(0.1464)	(0.2110)	(0.1705)	(1.3765)
AIDS	−0.0042*	0.6428	0.0101	−0.6995
	(0.0025)	(0.9711)	(0.1235)	(1.0167)
Adj.R^2	0.73	0.78	0.77	0.93
Obs.	39	19	25	7

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

OECD countries, for which our previous results show no effect of the epidemic on income levels for these countries, we expect the overall effect for the eighty-five countries to be insignificant. Contrary, for the non-OECD sample, we obtain a negative and significant at the 10% level estimate (column 3). The five outliers are part of the countries in Regime 1, obtained from the threshold estimation (Table 3). Since the non-OECD sample contains the bigger part of this group, the effect of the epidemic on income levels is much stronger. Figures 11 and 12 present the correlation between AIDS and GDP per worker in 2000 for the full sample (90 countries) and for the sample without the five potential outliers, correspondingly. What becomes apparent is that, AIDS is income-neutral. Both-high and low-income countries are affected by the epidemic.

4.4 AIDS and age groups

In addition to obtaining data on annual AIDS cases, we were also able to assemble data on the officially reported AIDS cases for the period of study on different age groups. We aggregate the data into four arbitrarily chosen age groups which we think are important when we study the

Table 5: Cross-country regressions for the three age groups

Dependent variable: ln(GDP per Worker)	
Specification	Extended Solow model with AIDS
Unrestricted	
Constant	5.2165***
	(1.0415)
ln s_{ik}	0.7308***
	(0.1474)
$\ln(n_i + g + \delta)$	−2.5760***
	(0.3181)
ln s_{ih}	0.4887***
	(0.0892)
AIDS(5-15)	−0.0024
	(0.0117)
AIDS(16-34)	−0.0056*
	(0.0030)
AIDS(35-60+)	0.0033
	(0.0042)
Adj.R^2	0.84
Obs.	63

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

impact of AIDS on income. The groups are: AIDS(0-4)-infancy period, AIDS(5-15)-schooling period, AIDS(16-34)-most productive period and AIDS(35-60+)-less productive period. Figures 8-10 present the distribution for selected countries in Africa, the Americas, and Europe for the four group ages. Some interesting observations become apparent from further disaggregating the AIDS data in different age groups. Two of the four groups - AIDS(16-34) and AIDS(35-60+) are affected most by the disease. In Africa, the most affected group is AIDS(16-34), when individuals are in their most productive stage of their lives. Similarly, in Europe this is the most affected group. Young people are more likely to be affected from HIV because they tend to have more sexual contacts. For the Americas, in countries like Argentina, Brazil and Mexico, the most affected group is AIDS(16-34) in comparison to the US in which the most affected group is AIDS(35-60+).

Similarly to obtaining AIDS incidence from the annual AIDS cases for each country, we obtain AIDS incidence for each one of the four age groups for all the countries, for which we were able to assemble data. Our sample (90 countries) was reduced to 63 countries.[17] We multiply the number of reported AIDS cases in each age group by 100,000 and divide by average population. In addition to the officially reported cases per age group, the WHO reported Not specified/unknown cases (NS). Although the data in the OECD countries does not or have very few NS cases, the data in most low-income countries like sub-Saharan countries contain a lot of NS cases, which are excluded from our regression. Data on population are taken for the WDI (2002). We start from the year, during which an AIDS case was reported till 2000. Contrary to obtaining AIDS incidence for the whole sample, where we start from the year during which a positive case was reported, here we start from the year, during which an AIDS case was reported.[18] Our dependent variable is the ln(GDP per worker in 2000). Data on the percentage of the working age population (15-64) are from the World Bank's *World Development Indicators* (2002). Our measure of human capital is taken from Bernanke and Gürkaynak (2001) and is the percentage of the working-age population that is in secondary school, averaged for the period 1970-1995. We average the population growth of the working-age population n, starting from the initial year during which an AIDS case was reported and add $g + \delta$, which is assumed to be 0.05. Following MRW, the saving rate s_k is the ratio of average investment to GDP over the same period as for AIDS (PWT 6.1). We present the results in Table 5.

Due to the high correlation between AIDS(0-4) and AIDS(16-34)-0.825, and AIDS(0-4) and AIDS(16-34)-0.812, we exclude AIDS(0-4) from our regression. AIDS(0-4) are the children of the older people and may get infected by HIV/AIDS only if their parents are HIV positive or they are already infected by AIDS. Our results show that the coefficient on AIDS(16-34) is negative and significant, implying negative impact of this AIDS infected age group on income level. The estimates on $\ln(s_{ik})$, $\ln(n_i + g + \delta)$, and $\ln(s_{ih})$ are all significant at the 1% level of significance.

[17]For details about the sample of countries and relevant variables used in the estimation, see the appendix, Table A2.

[18]The officially reported AIDS cases for the different age groups are reported before 1997 and annually for 1997, 1998, 1999 and 2000.

4.5 Panel-data estimation

In recent years there has been considerable empirical work on cross-country growth, explaining differences across countries. A common assumption in these studies is an identical production function for all the countries. For example, MRW assume a common Cobb-Douglas production function in which technology, A, grows exogenously at a rate g reflecting the advancement of knowledge, which is not country specific (i.e. $A(t) = A(0)e^{gt}$). Similarly, the rate of depreciation of capital is assumed to be constant across countries, so that $g + \delta = 0.05$. In contrast, $A(0)$ reflects not just technology, but also resource endowments, climate, institutions, and it may differ across countries. MRW assume that $\ln A(0) = a + \varepsilon$, where a is a constant and ε is country-specific shock assumed to be independent of savings, s, and population growth, n. This allowed them to estimate equation (1) with OLS. MRW provided three reasons for making this assumption of independence. First, this assumption is made not only in the Solow model, but also in many standard models of economic growth. In any model in which saving and population growth are endogenous but preferences are isoelastic, s and n are unaffected by ε. Second, although many economists have asserted that the Solow model cannot account for the international differences in income and this has stimulated work on endogenous-growth theory, the identifying assumption makes it possible to determine whether systematic examination of the data confirms these informal judgments. Third, the model predicts not just the signs but also the magnitudes of the coefficients on saving and population growth.

Islam (1995) questions the first assumption of independence – saving and population growth are unaffected by ε. In general, the country-specific technology term ε is likely to be correlated with saving and population growth rates experienced by that country. OLS is not valid and the estimated coefficients are potentially biased. One way to correct the endogeneity problem is to use instrumental variables. As Islam points out, it is difficult to come up with instruments that will be correlated with the included explanatory variables of the model and not correlated with A(0). Another solution to the endogeneity problem then is the use of panel data. By using the panel data technique, the unobserved heterogeneity in the initial level of efficiency is controlled (Temple, 1999). Despite these advantages, panel data techniques leave some uncertainty about the time intervals. Most researchers find it useful to use five or ten year averages to avoid business cycle effects.

This section extends our baseline cross-sectional results to consider estimation of the augmented Solow equation using panel data techniques. Even thought AIDS data for some countries exist since 1979, we consider the period $1985 - 2000$ since for most countries 1985 was the starting year of reported AIDS cases. This enables us to evaluate the impact of the epidemic across different countries and over time. The main advantage of panel data techniques is that they can allow for heterogeneity in the intercept, explaining the time variation for each country. Following Islam (1995), we average the data in five-year time intervals, since the disturbance terms are less likely to be influenced by business cycle fluctuations in this case. Thus, we have now three five-year time intervals-from $1985 - 1990$, $1990 - 1995$, and $1995 - 2000$.

Our regression equation is:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln A(0) + gt + \frac{\alpha}{1-\alpha-\beta}\ln\left(\frac{s_{it,k}}{n_{it}+g+\delta}\right) + \frac{\beta}{1-\alpha-\beta}\ln\left(\frac{s_{it,h}}{n_{it}+g+\delta}\right) + \varepsilon_{it}, \qquad (3)$$

where our dependent variable is the log GDP per working age person in country i at time t, n_{it} is the average population growth of the working age population, $g+\delta$ is assumed to be 0.05, and s_{it} is the ratio of average investment to GDP over the three five time-year periods.[19] Our measure of human capital is taken from Barro and Lee (2001) and is the percentage of secondary school attained in the total population. MRW use as a measure of human capital the percentage of the working-age population that is in secondary school. Our measure is different from theirs, but we don't expect this to lead to serious bias in the estimated coefficient on human capital. Similarly to the cross-country regressions, we add AIDS in the panel regressions.

Table 6 presents panel data analysis of the augmented Solow equation for the full sample under different specifications. First we estimate equation 3 using the between estimator. In a recent paper Hauk and Wacziarg (2004) argue that using an OLS estimator applied to a single cross-section of variables averaged over time (the between estimator) performs best in terms of the extent of bias on each of the estimated coefficients. The coefficient on AIDS is negative and significant at the 5% level, the estimated coefficients on $\ln(s_{it,k})$, $\ln(\mathrm{n}_{it}+g+\delta)$ and $\ln(\mathrm{s}_{it,h})$ have the expected signs and are significant at the 1% level (column 2).

To allow for the possibility of time effects, we have also estimated the model by adding $(T-1)$ time dummies[20], where $\mathrm{d91}_t$ and $\mathrm{d96}_t$ are dummy indicators for the years 1991 and 1996. This

[19] For detailed description of the data, refer to subsection 3.1 of the paper.

[20] In order to avoid perfect collinearity we drop the dummy variable on the first five-year period.

Table 6: Panel-data estimation

Dependent variable: ln(GDP per Worker(1985-2000))					
Specification	Extended Solow Model with AIDS				
Unrestricted	(*PWT* 6.1)				
	Non-oil BE	Non-oil with time effects	Non-oil FE with time and country effects	Non-oil with dOECD and inter. term	Non-oil with inter. term, dOECD, time effects
Constant	6.1223***	6.7409***	9.6703***	8.8542***	8.7658***
	(1.0323)	(0.5688)	(0.3080)	(0.6659)	(0.6395)
$\ln s_{it,k}$	0.6383***	0.6253***	−0.0664	0.5181***	0.5318***
	(0.1112)	(0.0003)	(0.0503)	(0.0663)	(0.0637)
$\ln(n_i + g + \delta)$	−1.8651***	−1.5647***	−0.0327**	−0.6969***	−0.6714***
	(0.3503)	(0.1932)	(0.1048)	(0.2500)	(0.2400)
$\ln s_{it,h}$	0.5464***	0.5532***	−0.3800***	0.5573***	0.5382***
	(0.0794)	(0.0512)	(0.0939)	(0.0496)	(0.0478)
AIDS_{it}	−0.0052**	−0.0047***	−0.0007	−0.0042***	−0.0048***
	(0.0022)	(0.0013)	(0.0006)	(0.0013)	(0.0012)
$d91$		0.1796			0.1641**
		(0.0724)			(0.0679)
$d96$		0.3196			0.3188***
		(0.0725)			(0.0676)
IT				0.0275**	0.0255**
				(0.0115)	(0.0111)
$dOECD$				0.4521***	0.4656***
				(0.1130)	(0.1085)
Adj. R^2	0.84	0.81	0.99	0.82	0.83
Obs.	82	241	241	241	241

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

captures exogenous shocks specific to each five-year period. The results are similar in terms of the magnitude and significance level to those obtained from estimating the model with the between estimator. There is a change in terms of the magnitude in the AIDS coefficient (-0.0047) and significance level - it is significant at the 1% level.

To account for the possibility of country specific effects as well as time effects, we estimate fixed effects (FE) with time and country effects. We notice a big change in terms of the magnitude of the coefficients and significance level. The estimate on $\ln(s_{it,k})$ becomes insignificant, the estimate on $\ln(\mathrm{s}_{it,h})$ changes from positive and significant in the previous two specifications into negative and significant. The estimate on $\ln(\mathrm{n}_{it}+g+\delta)$ is negative and significant at the 5% level of significance. The coefficient on AIDS is insignificant. These radical changes in some of the estimates is due to the loss of degrees of freedom., leading to insignificant estimates. Since this specification involves the addition of 79 ($N-1$) country specific dummy variables and 2 ($T-1$) time dummy variables, the sample size was reduced from 159 to 74.

In order to allow for the effect of AIDS to differ among OECD and non-OECD countries, we interact AIDS with an OECD dummy variable (IT hereafter). We report the results in column 5. All the estimates are significant and have the expected signs (column 5). The estimate on AIDS is negative and significant at the 5% level of significance. This is in agreement with our cross-country results that AIDS has a negative effect on income levels.

As a last robustness check we include in addition to the interaction term (the specification OLS with interaction term, column 5), time specific dummies (d91 and d96) to allow for the effect of AIDS to differ across time. The estimate on AIDS is negative and significant at the 10%, the IT is positive and significant and the dummy for OECD is positive. The estimates on $\ln(s_{it,k})$, $\ln(\mathrm{n}_{it}+g+\delta)$ and $\ln(\mathrm{s}_{it,h})$ are qualitatively similar to those obtained from the specification OLS with interaction term.

4.6 IV Estimation

In this section we present results from estimating the model using instrumental variables to correct for the endogeneity of AIDS.[21] As an epidemic that threatens economic performance of nations, AIDS is a function of factors like religion, culture, institutions, education and social infrastructure.

[21] We tried to instrument AIDS with initial AIDS in our cross-sectional analysis. Since initial AIDS is most likely to be measured with errors, this can bias our estimates toward zero.

Instrumental variables is one way to correct for endogeneity. It has been intensively used in empirical literature, but it has its own problems. We follow Woolridge (2002) and Ressler, Waters, Watson and Hill (2003). In a standard notation we write our main equation in the following way:

$$y_{it} = x_{it}^{'}\beta + Y_{it}^{'}\gamma + e_{it},$$

where

$$i = 1, ...N; t = 1, ..., T$$

Our dependent variable is y_{it} and is the income per working-age person for the period 1985-2000. The vector $x_{it}^{'}$ is (1xk) and includes the explanatory variables: $(s_{it,k})$, $(n_{it}+g+\delta)$ and ($\mathrm{s}_{it,h}$). The vector $Y_{it}^{'}$ is (1xm) and contains the possibly endogenous variable AIDS_{it}. It is difficult to come up with instruments that are correlated with the endogenous variables and not correlated with the error term. A common approach is the use of lags of the right-hand side variables as predetermined or weakly exogenous instruments in panel-data regressions. We use the first lag of AIDS and *schooling* as instrumental variables for AIDS. Past values of human capital play an important role in explaining the effect of AIDS on economic performance.

Instruments with no relevance for the endogenous variables are likely to give biased results. To prove the quality of our instruments, we test their validity by estimating reduced forms or regression of AIDS on the explanatory variables and the instrumental variables. Then we test the joint significance of the coefficients on the instruments in each of our specifications. In all the regressions, we reject the null hypothesis that the coefficients are zero at the .01 level of significance. This shows that our instruments provide useful information in addition to that provided by the explanatory variables.

A second test of the validity of our instruments is testing the overidentifying restrictions in each specification. We report the results in Table 7 (Panel B) with the results from IV estimation for the panel-data specifications (Panel A). For the specifications in column 2 and 3, the endogenous variable AIDS is explained with two instruments-lag of AIDS and lag of schooling. This results in one over-identifying restriction. For the next two specifications in column 4 and 5, in addition to AIDS we have one more potentially endogenous variable- the interaction term between AIDS and dummy OECD. Following Woolridge (2002), we include in our set of instruments an interaction term between dummy OECD and a lag of AIDS. This again results in one over-identifying restriction.

Table 7: IV Regressions

IV Regressions of ln(GDP per Worker)				
	Panel A: Two Stage Least Squares			
Specification				
Unrestricted	Non-oil with time effects	Non-oil FE with time and country effects	Non-oil with dOECD, and inter. term	Non-oil with iter. term, dOECD and time effects
Constant	7.1547***	10.3885***	9.3927***	9.3602***
	(0.7329)	(3.3262)	(0.8355)	(0.8316)
ln $s_{it,k}$	0.5698***	−0.3014	0.4922***	0.4921**
	(0.0891)	(0.2949)	(0.0853)	(0.0849)
$\ln(n_{it} + g + \delta)$	−1.4802***	−0.2226	−0.5320*	−0.5171*
	(0.2482)	(0.4997)	(0.3105)	(0.3091)
ln $s_{it,h}$	0.7290*	1.5568	0.5804***	0.5782***
	(0.1464)	(2.7382)	(0.0684)	(0.0680)
AIDS_{it}	−0.0082***	−0.0331	−0.0085***	−0.0090***
	(0.0021)	(0.0476)	(0.0020)	(0.0020)
d96	0.1322*	0.1547***		0.1512***
	(0.0778)	(0.0537)		(0.0735)
IT			0.0407**	0.0366**
			(0.0178)	(0.0174)
dOECD			0.4316***	0.4498
			(0.1538)	(0.1525)
Adj.R^2	0.78	0.50	0.81	0.81
Obs.	159	159	159	159
	Panel B: Specification Tests (p value)			
Overidetifying Restrictions	0.397	0.892	0.563	0.862
Hausman Test	0.204	0.993	0.088	0.077

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

In all the specifications we fail to reject the null of no correlation between the instruments and the error term. This shows that our over-identifying instruments are satisfactory. [22]

Now we use the Hausman test to determine whether AIDS should be treated as exogenous or endogenous. The results are reported in Panel B. In two of the specification - OLS with dOECD and interactive term (column 4) and OLS with dOECD, interaction term and time effects we are able to reject the null at the 10% level of significance that AIDS and the potentially endogenous interaction term are correlated with the error term. This implies that we can apply Two-Stage Least Squares and correct for endogeneity. For the specifications in columns 2 and 3 we are not able to reject the null.

The results from (2SLS) are presented in Panel A. The coefficient on $\ln(\mathrm{s}_{it,k})$ enters positive and significant at the 1% level, the coefficient on $\ln(\mathrm{n}_{it}+g+\delta)$ enters negative and highly significant and the estimate on $\ln(\mathrm{s}_{it,h})$ is positive, significant at the 10% level. The time dummy variable is positive and significant at the 10% level. The coefficient on AIDS is negative and highly significant, showing that even after we correct for endogeneity, we still obtain a negative and significant estimate. When we estimate IV regression on FE with time and country effects, we notice a big difference in terms of the magnitude and significance of the estimates. The estimates on $\ln(\mathrm{s}_{it,k})$, $\ln(\mathrm{n}_{it} + g + \delta)$, $\ln(\mathrm{s}_{it,h})$ become all insignificant. The estimate on AIDS also becomes insignificant. This might be due to the small time dimension of the panel and the large cross-sectional dimension, which leads to inaccurate estimate of a fixed effect for each country (Cook, 2000).

Column 4 presents results from OLS regression with dummy OECD and interaction term. The estimates have the expected sign and are significant. The estimate on AIDS is negative, significant at the 1% level, which is in accordance with our previous results. In column 5 (Panel A) we report results from IV regression on our last specification - OLS with interaction term, dummy OECD and time effects. The results are quantitatively similar to those from column 5. Only the estimate on dummy OECD changes from significant and positive into positive and insignificant.

5 Conclusion

In this paper, we investigate the impact of AIDS on income levels. Contrary to previous work on AIDS, we make use of the officially reported AIDS cases from WHO on 117 countries for the period

[22] We report the p-value from χ^2 Sargan's (1958) test. This is a test of the joint null hypothesis that the excluded instruments are valid instruments. A rejection casts doubt on the validity of the instruments.

1979-2000, during which AIDS as an epidemic has spread across the world. The emphasis is on level regressions in order to address the impact of AIDS on economic welfare measured by income.

First, we use the *extended Solow* model and we split the non-oil sample into OECD and non-OECD countries. In the full sample and non-OECD countries, the estimate of the coefficient on AIDS is negative. For the OECD countries, we obtain an insignificant coefficient estimate for AIDS, which shows that AIDS has no quantifiable effect on the income level for this sub-sample.

Then we use Hansen's threshold methodology to endogenously split the full sample. We show that AIDS is a threshold variable that splits the countries into four regimes, obeying different statistical models. When we estimate the four regimes, we find that AIDS has a negative impact on output per capita for the highly infected countries. In addition to the proposed in the literature threshold variables- output in 1960, literacy rate in 1960 and trade in 1985, we show that AIDS is a threshold variable.

Our panel-data analysis performs estimation of the augmented Solow model with AIDS under different specifications. AIDS enters negative and significant in all of them except the specification FE with country and time effects, where the estimate is insignificant. This might be due to the fact that fixed effects leave little additional variation to be explained. Our results are in agreement with the findings from the cross-section estimation for the period in question.

6 References

Acemoglu, D., S. Johnson and J. Robinson. (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* **91**, *1369-1401.*

Acemoglu, D., S. Johnson and J. Robinson (forthcoming). "Disease and Development in Historical Prospective," *Journal of the European Economic Association.*

Barro, R.J. (1991). "Economic Growth in a Cross Section of Countries," *The Quarterly Journal of Economics* **106**, 407-443.

Bernanke, B.S. and R.S. Gürkaynak (2001). "Is Growth Exogenous? Taking Mankiw, Romer and Weil seriously", *NBER Macroeconomics Annual* **16**, eds. Bernanke B. and K. Rogoff, MIT Press.

Bloom, D.E. and Ajay S. Mahal (1997). "Does the AIDS epidemic threaten economic growth?," *Journal of Econometrics* **77**, 105-124.

Bonnel, R. (2000). "HIV/AIDS: Does it increase or decrease growth in Africa?," Mimeo, World Bank.

Burger, R.(1996). "The Macroeconomic Impact of HIV/AIDS in South Africa:A Supply-side Analysis," University of Stellenbosch.

Cook, D. (2002). "Education and Growth: Instrumental Variables Estimates," *working paper, Hong Kong University of Science and Technology.*

Corrigan, P., G. Glomm and F. Mendez (forthcoming). "AIDS crisis and Growth," *Journal of Development Economics.*

Cuddington, J.T. (1993). "Modeling the Macroeconomic Effects of AIDS, with Application to Tanzania," *The World Bank Economic Review* **7(2)**, 173-189.

Cuddington, J.T. and J.D. Hancock (1994). "Assessing the impact of AIDS on the growth path of the Malawian economy," *Journal of Development Economics* **43**, 363-368.

Duffy, J. and C. Papageorgiou (2000). "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification"," *Journal of Economic Growth* **5**, 87-120.

Durlauf, S. and P. Johnson (1995). "Multiple Regimes and Cross–Country Growth Behavior," *Journal of Applied Econometrics* **10**, 365–84.

Easterly, W. and R. Levine (2001). "It's Not Factor Accumulation: Stylized Facts and Growth Models," *World Bank Economic Review* **13**, 129-158.

Ferreira, C. P, and S. Pessoa (2003). "The Long-Run Economic Impact of AIDS," Manuscript.

Frankel, J.A. and David Romer (1999). "Does Trade Cause Growth?," *American Economic Review* **89**, 379-399.

Gallup, J.L. and Jeffrey D. Sachs (2000). "The Economic Burden of Malaria," *CID Working Paper No.* **52**.

Hall, R. E. and C.I. Jones (1999). "Why Do Some Countries Produce so Much More Output Per Worker than Others," *Quarterly Journal of Economics* **114**, 83-116.

Hauk, W. and Jr. Roamin Wacziarg (2004). "A Monte Carlo Study of Growth Regressions," *NBER Technical Working Paper* no.**296**.

Hansen, B.E. (2000). "Sample Splitting and Threshold Estimation," *Econometrica* **68**, 575-603.

Islam, N. (1995). "Growth Empirics: A Panel Data Approach," *The Quarterly Journal of Economics* **110**, 1127-1170.

Kalemli-Ozcan, S. (2001). "The Effect of Mortality on Fertility and Human Capital Investment: What do we learn from AIDS?," *working paper,* University of Houston.

Kremer, M. (2002). "Pharmaceuticals and the Developing World," *Journal of Economic Perspectives* **16**, 67-90.

Levine, R., N. Loyaza and Th. Beck (2000). "Financial intermediation and growth: Causality and causes," *Journal of Monetary Economics* **46**, 31-77.

Mankiw, N.G., D. Romer and D.N. Weil (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* **107**, 407–37.

Masanjala, W. and C. Papageorgiou (2004). "The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity," *Journal of Applied Econometrics* **17**, 171-201.

Masanjala, W. and C. Papageorgiou (2004). "Rough and Lonely Road to Prosperity: A reexamination of the sources of growth in Africa using Bayesian Model Averaging," *working paper*, LSU.

McCarthy, F.Desmond, Holger Wolf and Yi Wu (2002). "The Growth Costs of Malaria," *working paper,* World Bank , Edmund A. Walsh School of Foreign Service and Georgetown University - Department of Economics.

Papageorgiou, C. (2002). "Trade as a Threshold Variable for Multiple Regimes," *Economics Letters* **77,** 85-91.

Papageorgiou, C., A. Savvides and M. Zachariadis (2004). "International Medical R&D Spillovers," *working paper,* Loisiana State University.

Temple, J. (1998). "Robustness Tests of the Augmented Solow Model," *Journal of Applied Econometrics* **13**, 361-375.

Temple, J. (1999). "The New growth Evidence," *Journal of Economic Literature* **37**, 112-156.

Ressler, R., M. S. Waters, J. K. Watson, and R. Carter Hill (2002). "Welfare Payments and the Spread of AIDS in the United States," *working paper,* Louisiana State University.

UNAIDS (2002), AIDS Epidemic Update.

UNAIDS/WHO Epidemiological Fact Sheets on HIV/AIDS and Sexually Transmitted Infections, 2002 update

Woolridge, J. M. (2002). "Econometric Analysis of cross section and panel data," *Cambridge, MA: MIT Press*

7 APPENDIX

The data

Data on GDP per capita and investment as a ratio to GDP were obtained from PWT (6.1).

Income (GDP per capita): Gross domestic product per capita in current prices (cgdp).

Investment : Investment share of GDP per capita in current prices.

Since data on real GDP per capita in 2000 are missing from the PWT 6.1, we use linear extrapolation for the following countries: Angola, Botswana, Central African Republic, Haiti, Mauritania, Papua New Guinea. Liberia and Somalia. Data on AIDS are not available for the following countries: Ivory Cost, Burma, Nepal, Zaire. We exclude Somalia due to the small number of observations -four, Germany was excluded from our sample. Sudan was excluded from our sample since data on all the variables from PWT 6.1 are available only for 1996.

AIDS Definition

In a meeting convened in Geneva by the WHO Global Programme on AIDS (1994) was suggested the following: the 1985 provisional WHO clinical case definition for AIDS ("Bangui definition") to be referred to as the WHO AIDS surveillance case definition and it was introduced an expanded WHO AIDS surveillance case definition. (Weekly Epidemiological Record,1994, 69, 273-280).

1. WHO case definition for AIDS surveillance

For the purposes of AIDS surveillance an adult or adolescent (>12 years of age) is considered to have AIDS if at least 2 of the following major signs are present in combination wit hat least 1 of the minor signs listed below, and if these signs are not known to be due to a condition unrelated to HIV infection.

Major signs

-weight loss >= 10% of body weight
-chronic diarrhoea for more than 1 month
-prolonged fever for more than 1 month (intermittent or constant)

Minor signs

- persistent cough for more than 1 month
- generalized pruritic dermatitis
- history of herpes zoster
- oropharyngeal candidiasis
- chronic progressive or disseminated herpes simplex infectiongeneralized lymphadenopathy

The presence of either generalized Kaposi sarcoma or cryptococcal meningitis is sufficient for the diagnosis of AIDS for surveillance purposes.

2. Expended WHO case definition for AIDS surveillance

For the purposes of AIDS surveillance an adult or adolescent (> 12 years of age) is considered to have AIDS if a test for HIV antibody gives a positive result, and 1 or more of the following conditions are present:

- >= 10% body weight loss or cachexia, with diarrhoea or fever, or both, intermittent or constant, for at least 1 month, not known to be due to a condition unrelated to HIV infection
- cryptococcal meningitis
- pulmonary or extra-pulmonary tuberculoses
- Kaposi sarcoma
- neurological impairment that is sufficient to prevent independent daily activities, not known to be due to a condition unrelated to HIV infection (for example, trauma or cerebrovascular accident)
- candidiasis of the oesophagus (which may be presumptively diagnosed based on the presence of oral candidiasis accompanied by dysphagia)
- clinically diagnosed life-threatening or recurrent episodes of pneumonia, with or without etiological confirmation
- invasive cervical cancer

Table A1: Data used in the extended Solow model

Country	Code	Values for relevant variables				
		$\frac{Y}{L}$	$\frac{I}{Y}$	SCHOOL	$n+g+\delta$	AIDS
Algeria	1	10005.4	13.65	0.0825	0.0829	0.1165
Angola	2	4360.1	6.35	0.0241	0.0773	3.5434
Argentina	3	18742.5	15.89	0.0859	0.0651	2.6654
Australia	4	40452.0	23.98	0.1108	0.0643	2.8723
Austria	5	36615.7	25.61	0.1075	0.0556	1.4293
Bangladesh	6	3046.7	10.30	0.0381	0.0733	0.0009
Belgium	7	38061.8	23.13	0.1094	0.0509	1.6902
Benin	8	2406.2	7.19	0.0252	0.0814	5.4167
Bolivia	9	5205.1	9.01	0.0646	0.0753	0.2169
Botswana	10	14769.7	17.38	0.0635	0.0816	57.0842
Brazil	11	11723.9	17.34	0.0587	0.0731	7.4395
Burkina Faso	12	2051.0	11.25	0.0073	0.0743	11.2315
Burundi	13	1248.1	6.07	0.0066	0.0698	27.4842
Cameroon	14	4321.1	6.64	0.0345	0.0785	10.8619
Canada	15	42080.2	24.97	0.1155	0.0619	3.0637
C.African Rep.	16	2357.0	5.11	0.0191	0.0723	20.3963
Chad	17	1903.4	6.63	0.0108	0.0760	12.7695
Chile	18	16137.4	18.79	0.0941	0.0673	1.7143
Colombia	19	9276.3	12.14	0.0834	0.0748	1.5264
Congo	20	5024.4	7.48	0.1059	0.0787	168.5997
Costa Rica	21	9391.8	16.04	0.0806	0.0789	3.4051
Denmark	22	42759.9	22.52	0.1151	0.0533	2.4675
Dom. Repub.	23	9089.1	13.43	0.0764	0.0744	4.2897
Ecuador	24	6051.4	15.90	0.0917	0.0798	0.7835
Egypt	25	7282.9	6.06	0.1082	0.0759	0.0295
El Salvador	26	7778.1	7.85	0.0525	0.0754	3.2685
Ethiopia	27	1388.1	4.27	0.0179	0.0743	7.1639
Finland	28	36433.6	24.42	0.1164	0.0526	0.3876
France	29	36165.8	24.60	0.1065	0.0550	4.8720
Ghana	30	2464.5	6.08	0.0678	0.0841	16.6795
Greece	31	23087.6	21.53	0.0968	0.0563	1.2263
Guatemala	32	8202.7	7.40	0.0350	0.0782	2.2228
Haiti	33	6235.0	5.31	0.0256	0.0228	8.1973
Honduras	34	3947.2	14.48	0.05034	0.0334	13.2563
Hong Kong	35	38179.1	25.05	0.0859	0.0215	0.4939
India	36	4360.6	12.35	0.0609	0.0217	0.0734
Indonesia	37	6263.5	17.76	0.0629	0.0237	0.0159
Ireland	38	40520.7	19.79	0.1453	0.0623	1.0947
Israel	39	30942.5	26.60	0.1163	0.0804	0.8832
Italy	40	33816.6	22.27	0.0836	0.0523	4.5305
Jamaica	41	5648.5	17.72	0.1233	0.0667	11.1127
Japan	42	38057.5	32.56	0.1038	0.0532	0.0950
Jordan	43	7490.8	15.15	0.1548	0.1016	0.1469
Kenya	44	2451.1	8.07	0.0417	0.0868	24.9535
Korea	45	20719.5	36.29	0.1261	0.0653	0.0306
Madagascar	46	1677.6	3.03	0.0383	0.0789	0.0211

Table A1: Data used in the extended Solow model						
Country	Code	Values for relevant variables				
		$\frac{Y}{L}$	$\frac{I}{Y}$	SCHOOL	$n+g+\delta$	AIDS
Malawi	47	1591.9	7.92	0.0147	0.0740	40.9708
Malaysia	48	15251.6	26.56	0.0906	0.0794	1.6425
Mali	49	1995.9	8.23	0.0162	0.0747	3.7066
Mauritania	50	2984.3	8.70	0.0201	0.0797	2.0821
Mauritius	51	21132.0	12.52	0.0808	0.0654	0.4024
Mexico	52	15629.6	17.49	0.0953	0.2901	2.9271
Morocco	53	7024.9	11.95	0.0547	0.0767	0.2073
Mozambique	54	2107.5	3.41	0.0112	0.0686	9.8234
Netherlands	55	37847.2	22.58	0.1226	0.0569	1.8466
New Zealand	56	30608.2	22.20	0.1223	0.0613	1.1704
Nicaragua	57	3584.3	12.41	0.0775	0.0832	0.4314
Niger	58	1875.0	4.61	0.0091	0.0837	4.2395
Nigeria	59	1592.5	9.39	0.0330	0.0793	3.1480
Norway	60	49423.1	28.65	0.1129	0.0559	0.9070
Pakistan	61	3956.5	11.14	0.0359	0.0745	0.0112
Panama	62	10528.0	18.78	0.1079	0.0749	7.7935
Papua N.G.	63	5778.8	10.35	0.0218	0.0785	1.5274
Paraguay	64	8423.9	12.70	0.0558	0.0818	0.6948
Peru	65	7767.1	17.62	0.1068	0.0753	2.3352
Philippines	66	6896.7	14.36	0.1239	0.0771	0.0420
Portugal	67	25241.1	23.10	0.0836	0.0532	4.8888
Rwanda	68	1839.0	4.64	0.0101	0.0785	18.5401
Senegal	69	3161.3	6.71	0.0258	0.0789	2.5547
Sierra Leone	70	1388.0	4.85	0.0258	0.0718	0.5959
Singapore	71	40393.7	42.45	0.0971	0.0759	1.3665
South Africa	72	12844.5	9.01	0.0881	0.0779	2.2471
Spain	73	27861.2	24.47	0.1157	0.0561	8.4116
Sri Lanka	74	5695.3	12.34	0.1030	0.0698	0.0467
Sweden	75	38254.8	21.12	0.0960	0.0536	1.1200
Switzerland	76	41885.1	27.79	0.0946	0.0565	5.6556
Syrian Arab Republic	77	7742.7	9.17	0.1052	0.0891	0.0360
Tanzania	78	932.4	16.46	0.0079	0.0822	26.0605
Thailand	79	9858.3	32.98	0.0570	0.0695	17.0469
Togo	80	1760.4	8.12	0.0425	0.0798	21.9104
Trinidad and Tobago	81	20072.5	9.39	0.1175	0.0650	21.9104
Tunisia	82	11064.1	13.26	0.0695	0.0770	0.4423
Turkey	83	11548.5	18.80	0.0740	0.0726	0.0376
Uganda	84	2132.7	13.65	0.0172	0.1196	19.1190
United Kingdom	85	37153.1	18.77	0.0998	0.1283	1.6040
Uruguay	86	16503.9	10.76	0.0907	0.1240	2.8308
USA	87	53979.1	21.29	0.1163	0.1198	14.8092
Venezuela	88	11757.8	14.30	0.0686	0.1043	2.6470
Zambia	89	1664.6	8.94	0.0367	0.0972	39.7673
Zimbabwe	90	5053.0	13.49	0.0577	0.1077	55.4721

Table A2: Data used in the extended Solow model with age groups					
Country	Code	Values for relevant variables			
		AIDS(0-4)	AIDS(5-15)	AIDS(16-34)	AIDS(35-60)
Algeria	1	0.0122	0.0448	0.8715	0.8798
Angola	2	0.9356	0.8842	18.3926	12.6764
Argentina	3	2.8651	0.6549	29.6146	14.3829
Australia	4	0.1493	0.1321	18.7920	29.9765
Austria	5	0.3192	0.0894	6.0399	18.7583
Belgium	6	0.8683	0.4192	10.2801	15.5798
Bolivia	7	0.0329	0.0060	0.4577	0.3619
Brazil	8	3.9651	1.0459	78.0027	57.3414
Burkina Faso	9	1.7485	0.6769	30.7855	7.2085
Canada	10	0.55542	0.1815	24.7142	38.4108
Chad	11	4.6754	1.3853	32.2871	70.3043
Chile	12	0.3989	0.1227	15.1747	16.3101
Colombia	13	0.4511	0.0833	11.8716	7.7976
Costa Rica	14	0.9189	0.4753	32.3824	28.8653
Denmark	15	0.2506	0.0771	16.0201	28.1075
Dom. Repub.	16	1.7630	0.6611	42.7656	42.9722
Ecuador	17	0.1023	0.0837	5.9734	4.8196
Egypt	18	0	0.0154	0.2104	0.3127
El Salvador	19	2.4291	0.4265	29.1126	19.8410
Finland	20	0.0399	0.0199	2.2929	3.9078
France	21	1.0198	0.4245	42.8421	49.2251
Ghana	22	4.5661	1.9071	143.3103	115.1556
Greece	23	0.2838	0.1860	8.2700	10.4035
Guatemala	24	1.5028	0.3542	24.3236	13.7397
Honduras	25	9.9932	2.8688	122.4458	70.2187
Hong Kong	26	0.0837	0.0670	2.8463	5.3745
Indonesia	27	0.0028	0.0017	0.1519	0.0920
Ireland	28	0.4746	0.3350	11.5578	7.2027
Israel	29	0.3203	0.1001	5.6246	7.2459
Italy	30	0.8304	0.3493	47.383	33.2371
Jamaica	31	12.6922	2.6373	78.2137	81.7577
Japan	32	0	0.0113	0.4731	0.9189
Jordan	33	0.0524	0.1049	1.0225	1.1274
Korea	34	0	0	0.1194	0.3245
Madagascar	35	0	0	0.1201	0.1801
Mauritius	36	0.1811	0	2.6259	2.8976
Mexico	37	0.8553	0.5060	28.7033	24.8432
Morocco	38	0.0840	0.0168	2.4541	0.7732
Netherlands	39	0.2057	0.1327	12.8610	22.2844
New Zealand	40	0.1423	0.0854	7.9698	12.5240
Nicaragua	41	0.0469	0.0234	3.9130	2.3900
Niger	42	0.7993	0.1827	23.26031	32.5096
Norway	43	0.0701	0.1402	6.0763	10.1661
Pakistan	44	0.0043	0.0017	0.0830	0.0770
Panama	45	5.5023	1.7672	69.4011	80.3655

Table A1: Data used in the extended Solow model					
Country	Code	Values for relevant variables			
		AIDS(0-4)	AIDS(5-15)	AIDS(16-34)	AIDS(35-60)
Papua N.G.	46	0	0	0.2286	0.3886
Paraguay	47	0.3310	0.0662	6.3550	3.5306
Peru	48	2.6048	0.2754	24.3510	17.3716
Portugal	49	0.6434	0.3318	43.7719	33.0651
Singapore	50	0.1789	0	7.4823	15.1732
Spain	51	1.8836	0.6227	95.1679	54.403
Sri Lanka	52	0.0118	0.0059	0.1836	0.4855
Sweden	53	0.1976	0.0581	9.5996	10.4828
Switzerland	54	1.1310	0.3231	51.6434	46.4732
Syrian Arab Rep.	55	0.0150	0.0224	0.2392	0.2841
Tanzania	56	16.2243	5.3510	220.3146	118.6307
Thailand	57	11.9101	26.1015	197.1519	93.9816
Togo	58	22.5089	1.9026	135.4046	181.7396
Tri.&Tobago	59	14.7051	2.1123	116.2592	103.3415
Turkey	60	0.0051	0.0103	0.2666	0.4060
United Kingdom	61	0.5480	0.2126	13.1307	16.4255
Uruguay	62	1.7178	0.1591	25.4496	16.0969
USA	63	2.6757	0.7804	114.2215	163.9776

Figure 5: GDP per capita vs. AIDS in Africa

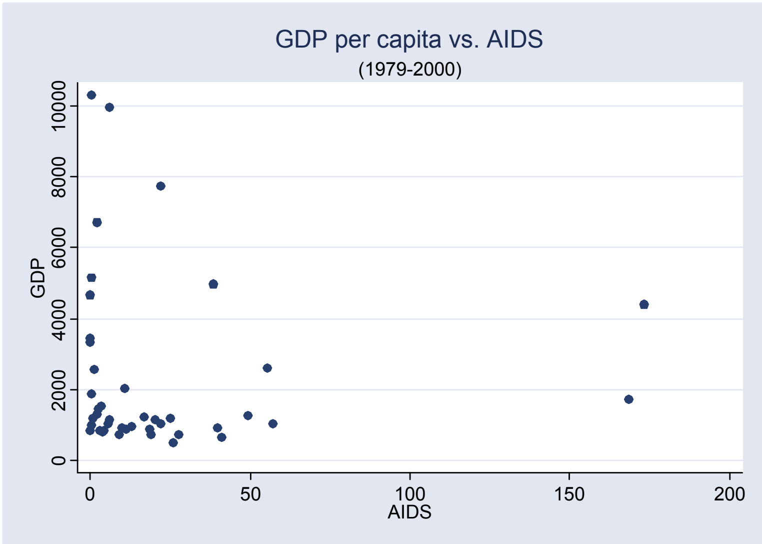


Figure 6: GDP per capita vs. AIDS in Europe

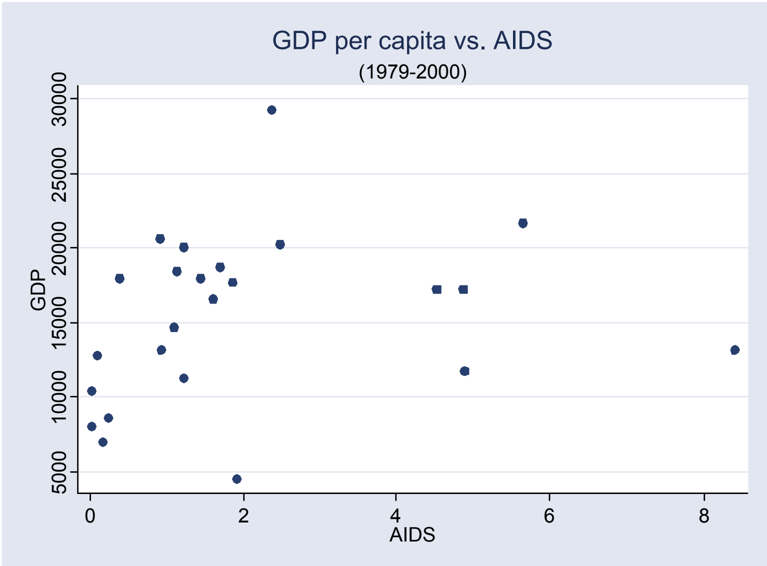


Table 8: Cross-country regressions for the full sample and for OECD and non-OECD countries

Dependent variable: Growth GDP per Worker (initial-2000)						
Specification	Extended Solow Model *(PWT 6.1)*			Extended Solow Model with AIDS *(PWT6.1)*		
	Non-oil	OECD	Non-OECD	Non-oil	OECD	Non-OECD
Unrestricted						
Constant	4.1891 ***	4.2603***	5.3233 **	4.2436*	3.9329**	5.4441**
	(2.2327)	(1.6681)	(2.5389)	(2.2835)	(1.6943)	(2.6062)
ln y_i	−0.3028	−0.1487	−0.4237**	−0.3072	−0.1567	−0.4333**
	(0.1924)	(0.1290)	(0.2130)	(0.1969)	(0.1284)	(0.2187)
$\ln s_{ik}$	0.5314***	0.0258	0.4777***	0.5301***	−0.0417	0.4743***
	(0.1810)	(0.2233)	(0.1552)	(0.1811)	(0.1997)	(0.1548)
$\ln(n_i + g + \delta)$	−0.2511	0.2197 **	−0.1801	−0.2485	0.1242	−0.1792
	(0.1670)	(0.0906)	(0.2013)	(0.1662)	(0.0935)	(0.3180)
$\ln s_{ih}$	0.2382***	0.6071**	0.2789*	0.2408*	0.6081 **	0.2845*
	(0.1392)	(0.2515)	(0.1435)	(0.1423)	(0.2588)	(0.0941)
AIDS				−0.0006	0.0138*	−0.0011
				(0.0013)	(0.0076)	(0.0015)
Adj. R^2	0.36	0.32	0.34	0.36	0.38	0.34
Obs.	90	21	69	90	21	69

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level. The investment and population growth rates are averages for the period 1979-2000. s_h is the average percentage of the working-age population in secondary school for the period 1970-1995.

Figure 7: Regression Tree Diagram

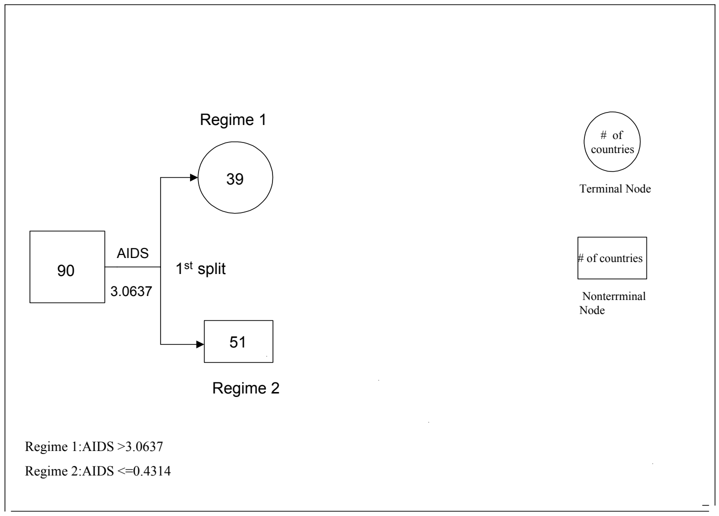


Table 9: Countries from the two regimes

Regime1		Regime 2		
Angola	Malawi	Algeria	Japan	Sri Lanka
Benin	Mali	Argentina	Jordan	S. Africa
Botswana	Mozambique	Australia	Korea	Syrian Arab Rep.
Brazil	Niger	Austria	Madagascar	Sweden
Burkina Faso	Nigeria	Bangladesh	Malaysia	Turkey
Burundi	Panama	Belgium	Mauritania	Tunisia
Cameroon	Portugal	Bolivia	Mauritius	U.K.
C. Africa	Rwanda	Canada	Mexico	Uruguay
Chad	Spain	Chile	Morocco	Venezuela
Congo	Switzerland	Colombia	Netherlands	
Costa Rica	Tanzania	Denmark	N. Zealand	
Dom.Rep.	Thailand	Ecuador	Nicaragua	
El Salvador	Togo	Egypt	Norway	
Ethiopia	Tr.&Tobago	Finland	Pakistan	
France	Uganda	Greece	Papua N.G.	
Ghana	USA	Guatemala	Paraguya	
Haiti	Zambia	Hong Kong	Peru	
Honduras	Zimbabwe	India	Philippines	
Italy		Indonesia	Senegal	
Jamaica		Ireland	Sierra Leone	
Kenya		Israel	Singapore	
(39)		(51)		

Table 10: Cross-country regressions for the two regimes

Dependent variable: ln(GDP per Worker in 2000)		
Specification	Regime 1	Regime 2
Unrestricted		
Constant	9.3328***	11.1989***
	(2.7346)	(0.7285)
ln s_{ik}	0.5474**	0.9397***
	(0.2023)	(0.1444)
$\ln(n_i + g + \delta)$	−1.1220	−0.5458**
	(0.8748)	(0.2102)
ln s_{ih}	0.7290*	0.6713***
	(0.1464)	(0.1442)
AIDS	−0.0042*	0.2724***
	(0.0025)	(0.0573)
Adj.R^2	0.73	0.80
Obs.	39	51

Notes: Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

Table 11: Robustness to potential outliers for the Non-Oil and Non-OECD sample

Dependent variable: ln(GDP per Worker in 2000)		
Specification	Extended Solow Model with AIDS *(PWT 6.1)*	
	Non-oil	Non-OECD
<u>*Unrestricted*</u>		
Constant	11.3786***	11.5750***
	(1.0085)	(0.7507)
$\ln s_{ik}$	0.8505***	0.6069***
	(0.1366)	(0.1489)
$\ln(n_i + g + \delta)$	-0.4910	-0.0466
	(0.3238)	(0.2936)
$\ln s_{ih}$	0.6223***	0.5232***
	(0.1219)	(0.1193)
AIDS	-0.0126	-0.0193*
	(0.0122)	(0.0109)
Adj. R^2	76	68
Obs.	85	64

Notes: The Non-oil sample includes 85 countries after we exclude Botswana, Congo, Malawi, Zambia, Zimbabwe. Standard errors are in parentheses. It is assumed that $g + \delta = 0.05$ as in MRW. All regressions are estimated using OLS. White's heteroskedasticity correction was used. *** Significantly different from 0 at the 1% level. ** Significantly different from 0 at the 5% level. * Significantly different from 0 at the 10% level.

Figure 8: Number of AIDS cases per age group in selected African countries

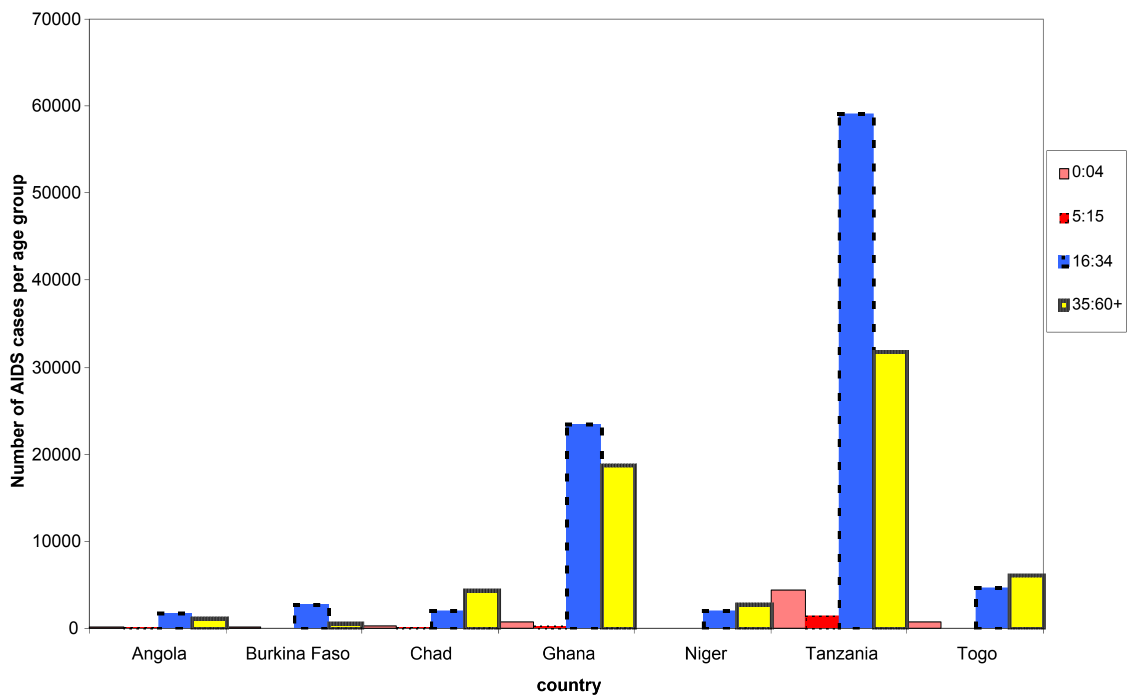


Figure 9: Number of AIDS cases age group in selected countries in the Americas

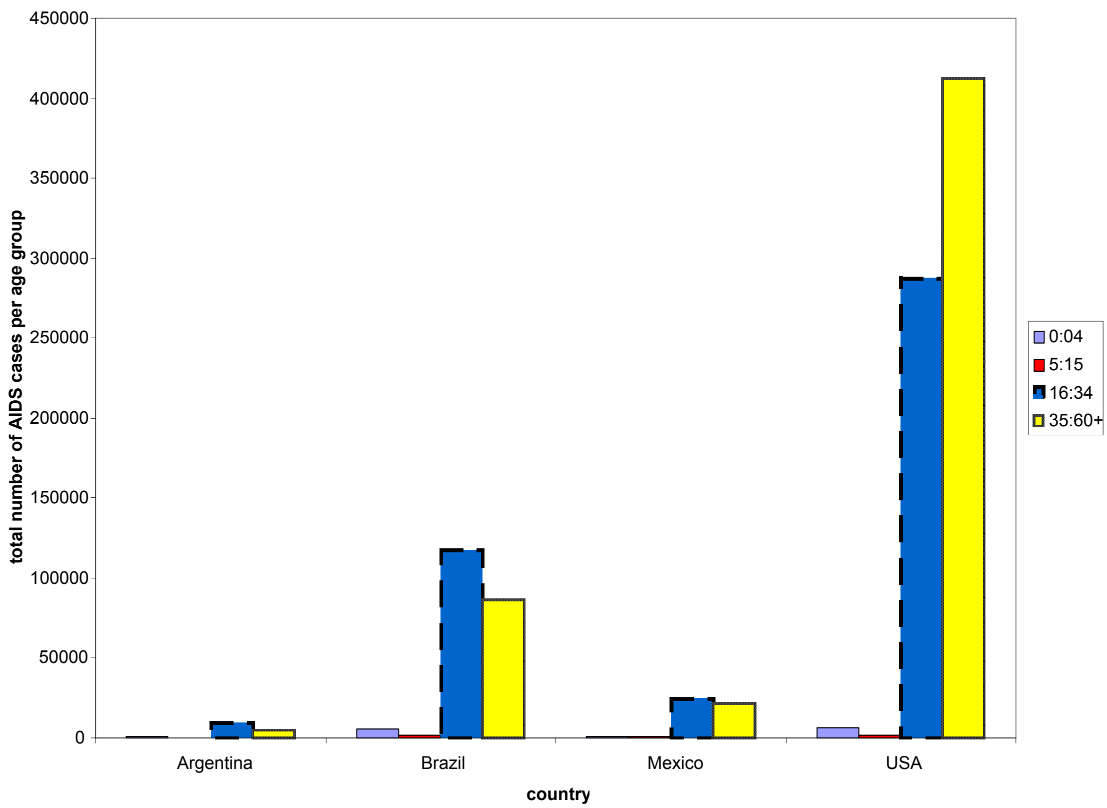


Figure 10: Number of AIDS cases in selected European countries

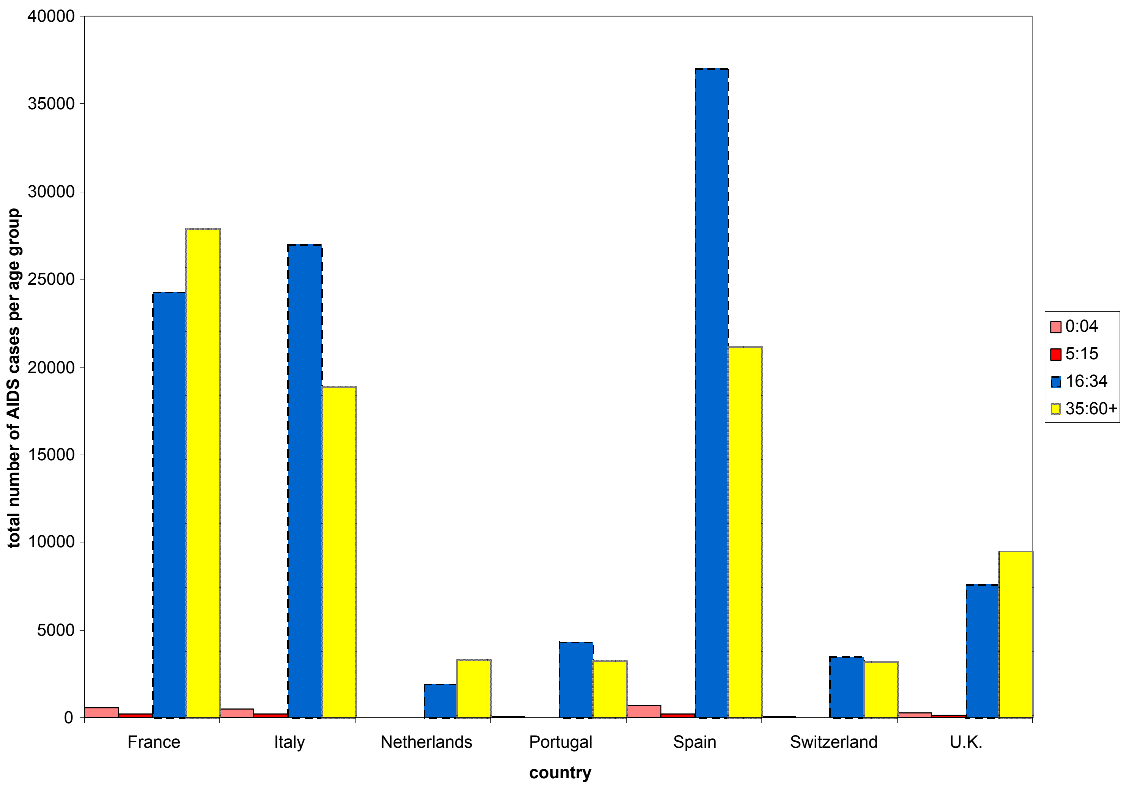


Figure 11: Income vs. AIDS (1979-2000)

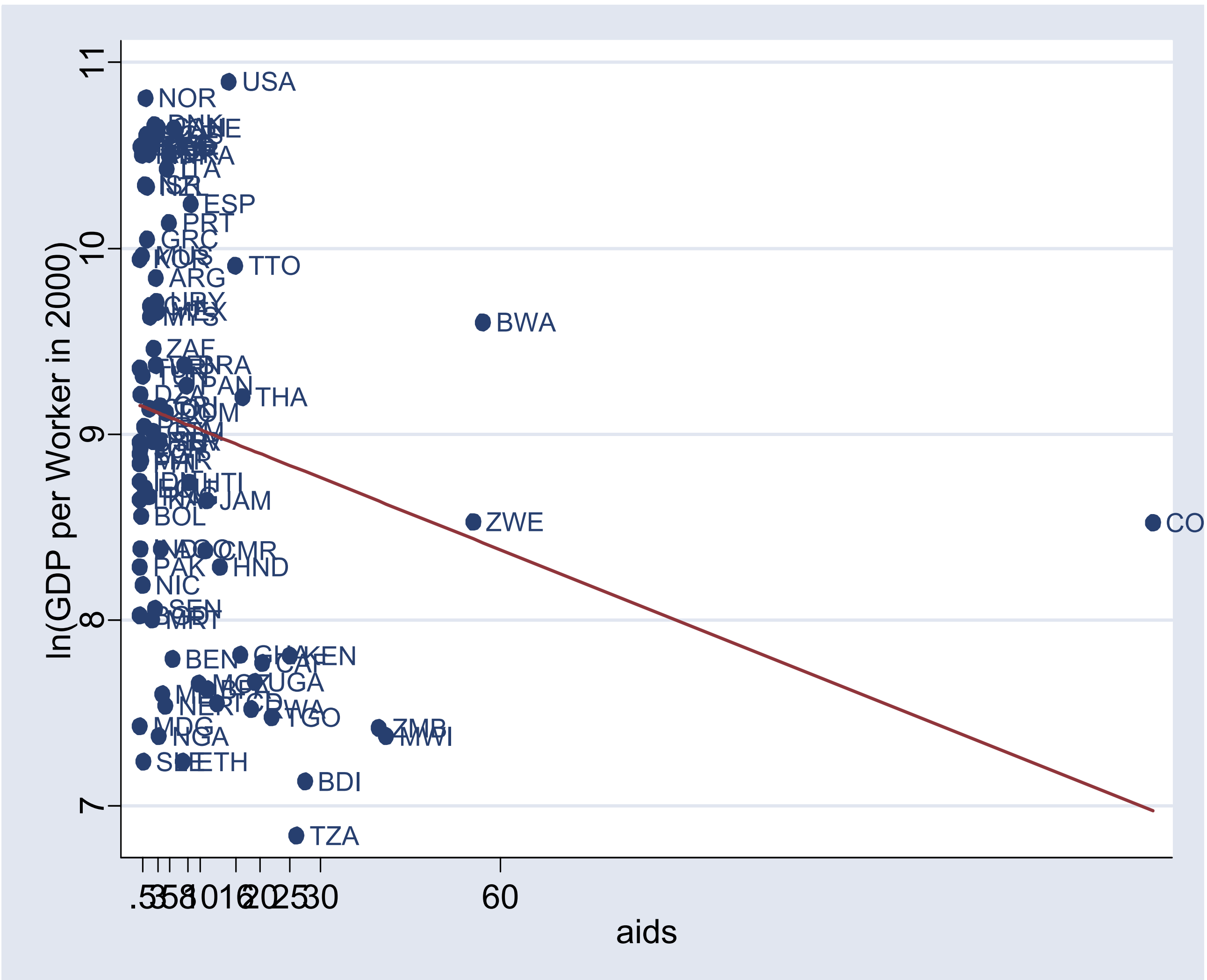


Figure 12: Income vs. AIDS (1979-2000) without potential outliers