UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Life expectancy and economic growth: AIDS as a natural experiment

Jessica Wirström
Andreas Bengtsson

Abstract
The purpose of this study is to examine how life expectancy, as a measure for health, affects economic growth in Sub-Saharan Africa. Furthermore, the model specifications in this thesis include AIDS, which is assumed to have an exogenous effect on life expectancy. This effect is in turn regarded as a natural experiment for life expectancy. An econometric model, based on the Solow growth model and previous research, is constructed. The data covers 16 sample countries across Sub-Saharan Africa between the years 1990-2017. Life expectancy at birth is the variable of interest, proxying the effect of health, which is presumed to be exogenously affected by AIDS-related deaths. AIDS-related deaths is also included as a variable of interest, to investigate if the epidemic alone affected economic growth, and not solely through its assumed effect on life expectancy. Additional control variables, which are all related to economic growth, are included to minimize omitted variable bias. Results from the model show that there is indeed a statistically significant relationship between life expectancy at birth and GDP growth. However, the findings show that inference regarding the direct effect of AIDS-related deaths on the dependent variable could not be drawn, as it was estimated as insignificant. The study concludes that the reduced life expectancy, seen during the AIDS epidemic, has had a negative impact on GDP growth in Sub-Saharan Africa. Given the assumptions about AIDS exogenous effect on life expectancy, this would imply that without the negative effect of AIDS on life expectancy, the economic growth in Sub-Saharan Africa would not have been as impeded.

Bachelor's thesis (15hp)
Department of Economics
School of Business, Economics and Law
University of Gothenburg
Supervisor: Dick Durevall

Acknowledgements: We would like to extend our gratification to our supervisor Prof. Dick Durevall.

List of abbreviations

ARD	**AIDS-related deaths**
GDP	**Gross Domestic Product**
IMF	**International monetary fund**
FCF	**Fixed Capital Formation**
ART	**Antiretroviral treatment**
WNEPI	**Weighted net export price index**

Contents

1. Introduction

Health is acknowledged as being a key feature for human- and economic development (World Health Organization, 2019). Public health is commonly found to be poor in low developed countries, which includes major parts of Africa. In 2016, the second largest cause of death in Africa was AIDS, after lower respiratory infections (WHO, 2016). Even though both of these causes are health related, an important difference to emphasize is that the latter mainly affects children and elderly, whereas HIV/AIDS affects the entire population, including the working-age cohorts.

Many countries in Sub-Saharan Africa struggle greatly with HIV/AIDS, it is the region where the largest number of people infected by HIV live. From 1990 and a few years into the 21st century, when the HIV/AIDS epidemic was at its peak, life expectancy at birth decreased with roughly 20% in countries such as Botswana, Eswatini and South Africa (World Development Indicators, 2019). The number of HIV infected people relative to AIDS victims continues to increase, a consequence of the improved access to antiretroviral treatment (ART) which both improves lives and has caused a grand reduction in the number of AIDS-related deaths (Avert, 2017). The access to antiretroviral treatment is also expected to be the main reason for the increase in life expectancy in the 21st century. In conclusion, during the most severe years of AIDS, life expectancy decreased. Later, when ART was introduced, life expectancy increased. An exogenous effect of AIDS on life expectancy is therefore assumed in this thesis which makes the effect of AIDS-related deaths on life expectancy an applicable natural experiment, further discussion regarding the exogeneity of the effect is found in section 2.2.

Sub-Saharan Africa is a region characterized by low economic development, corruption and conflicts, making it sensitive to shocks, such as epidemics. By using the assumed exogenous effect that the AIDS epidemic has had on life expectancy at birth, this thesis aims to study if life expectancy has affected economic growth in 16 selected Sub-Saharan countries (table 1).

One study analyzing the relationship between health and economic growth is written by Lorentzen, McMillian and Wacziarg (2005). They looked at how adult mortality affects economic growth of income per capita and found that adult mortality changes behaviour in ways which disfavour growth. Another study regarding health by Acemoglu and Johnson (2007) investigated the effect health improvements has had on economic growth. Due to

omitted disadvantages the authors claimed that causality between health improvements and growth cannot be established. The authors did nonetheless state that since AIDS affects productive adults it is likely to have a negative effect on growth. However, studies specifically examining the effect AIDS has had on economic growth show varying results. For instance, Over (1992) estimated a comparison between scenarios with and without AIDS from 1990 to 2040 to investigate if the epidemic would have a macroeconomic effect. Over aimed to find out if the epidemic affected GDP per capita in other aspects than population growth rate reduction, which alone would increase GDP per capita. The author concluded that AIDS has had a substantial negative impact on economic growth.

> The net effect of the AIDS epidemic on the growth of per capita GDP is a reduction of about a third of a percentage point in the ten countries with the most advanced epidemics. This is a substantial impact in countries that have been struggling to escape from a period of negative growth rates. (Over, 1992, p. 1).

In strong contrast, Alwyn Young (2005) studied the effect AIDS would have on the future standards of living. The author found that due to the decreased population growth rate the region would be left with a relative redundancy of resources, which would increase future consumption and thereby affect economic growth positively.

Although several studies looking at the relationship between economic growth and HIV/AIDS have been conducted through the years, most of these have focused on either a relatively concise time frame or on a rather limited country subset. In addition, the studies have generally regarded the sum of HIV-infected people as the variable of interest and have chosen control variables which are mainly correlated with the effectiveness of health care, e.g. human capital measures (Dixon, McDonald, & Roberts, 2001; Kahende & Hoch, 2008; Bloom & Mahal, 1997). This thesis on the other hand investigates the economic impact of AIDS-related deaths relative to the labour force and controls for variables assumedly associated with Gross Domestic Product (GDP) growth.

Numerous studies regarding the relationship between AIDS and GDP growth were published during the most severe time of the epidemic, from 1990 until early 2000. The introduction of antiretroviral treatment in 1997 changed the situation dramatically by reducing the number of AIDS-related deaths (Avert, 2017). The impact of the antiretroviral drug was deferred and the

effect on life expectancy can first be observed in the 21st century. Investigating the relationship between AIDS and GDP growth today, in 2019, would therefore probably give different results since it is now possible to observe the increase, peak and decline of the epidemic, including the effect from the introduction of ART.

The purpose of this thesis is to examine how life expectancy at birth, as a measure for health, has affected economic growth in Sub-Saharan Africa. The assumed exogenous effect of AIDS on life expectancy is used as a natural experiment. We conduct a panel data analysis with the Solow growth model as a foundation to achieve the purpose of the thesis. To reach a conclusion the following research question will be answered.

- Does life expectancy, as a measure for health, have an impact on economic growth in Sub-Saharan Africa?

The thesis proceeds as following; Firstly, a background of HIV/AIDS, an introduction to the economic theory used and a brief literature review. Secondly, a description of the data and the method used to analyse the data. Thereafter, results of the regressions and econometric model and a sensitivity analysis are presented, followed by a discussion including an analysis of the results. Lastly, the conclusion, answering the research question is presented, followed by the references and appendix.

2. Background

2.1 Background on HIV/AIDS

HIV was first documented as an upcoming epidemic in the late 1970's and by the 1980's HIV was estimated to have spread to five continents. The following decade, in the 1990's, the virus quickly advanced and was found to be the cause of AIDS (Avert, 2017). The virus increased rapidly worldwide until the peak year of 1997. After years of research on a treatment for the virus, 1996 marked the beginning of the era of Highly Active Antiretroviral Therapy. A year later, in 1997, the first combination antiretroviral pill was approved, which massively improved HIV treatment (Avert, n.d.). This is the course of the HIV/AIDS epidemic which will be analysed in the thesis, starting when the epidemic escalated in 1990, increasing until its peak year of 1997 and thereafter declining due to antiretroviral treatment.

In Sub-Saharan Africa, the virus hit catastrophically, and today in 2019 it is still a highly topical issue globally. Sub-Saharan Africa is home to approximately 50% of the total number of people infected by HIV in the world. The number of HIV infected people is continuously increasing in Sub-Saharan Africa, although the yearly number of AIDS-related deaths (ARD) has decreased from 760,000 in 2010 to 420,000 in 2016 (Avert, 2017).

Figure 1 illustrates the AIDS-related deaths for each of the sample countries from 1990-2017, as a share of labour in percent. For most of the countries, the figure shows the escalation of AIDS-related deaths until the beginning of the 21st century, followed by the subsequent decline as described above.

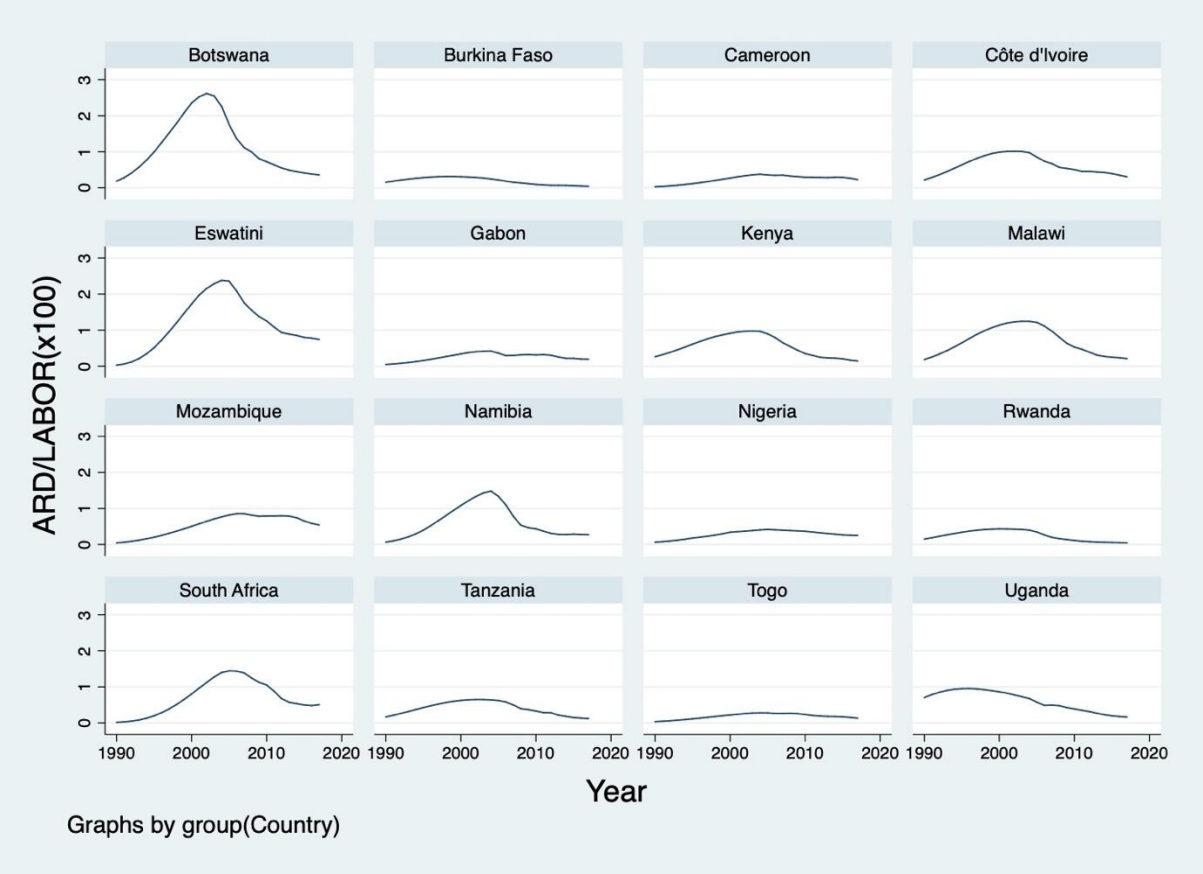

Figure 1: AIDS-related deaths/Labour in %, 1990-2017.

HIV/AIDS has certainly had a negative social impact on the countries severely affected. There is however also some evidence that the deaths that followed the epidemic decreased life expectancy at birth (appendix table A2), which consequently is expected to have had an impact on GDP growth. Figure 2 illustrates how life expectancy at birth changed in each of the sampled

Sub-Saharan countries during the studied time period. The figure shows a decline in life expectancy for several of the sample countries from 1990 until the beginning of the 21st century, after that, it starts to increase.

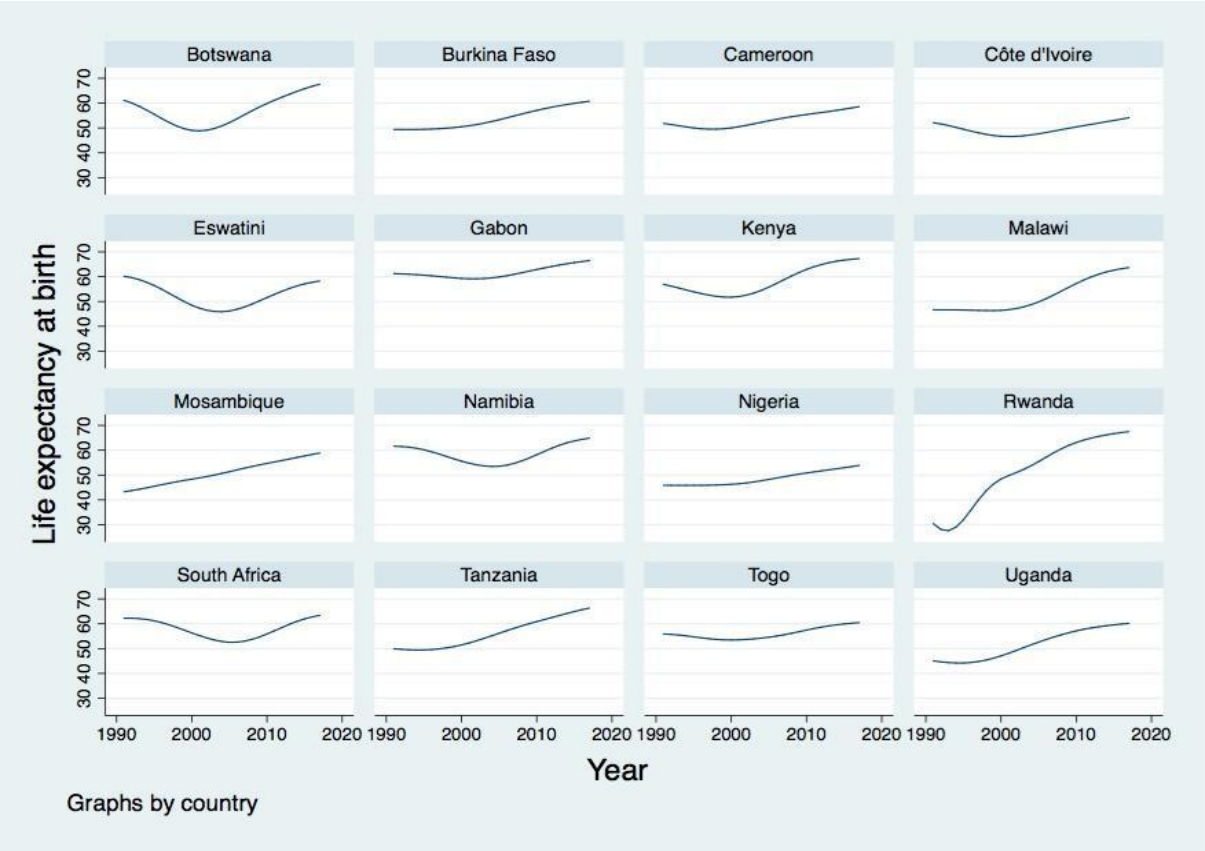

Figure 2: Life expectancy at birth (1990-2017)

In Botswana during the worst years of the AIDS epidemic, life expectancy at birth declined from nearly 62 years in 1990, to 49 years in 2001. It then increased, probably in part due to the introduction of antiretroviral treatment, reaching 67 years in 2017. In Uganda, the decline in life expectancy during the studied time period was not as severe as in Botswana. The country did however experience a decline from 45,5 years to 44,2 years between 1990 and 1995, and then an escalation to 60 years in 2017 (World bank, 2019).

Figure 3 shows the negative relationship between life expectancy at birth (vertical axis) and AIDS-related deaths (horizontal axis) in 6 of the sample countries. There is a negative correlation for all the countries seen in figure 3, though not in equal magnitude, as also explained when comparing Botswana and Uganda. The relationship provides some evidence

that the assumed exogenous effect of ARD is an appropriate natural experiment to estimate how changes in life expectancy at birth affect economic growth.

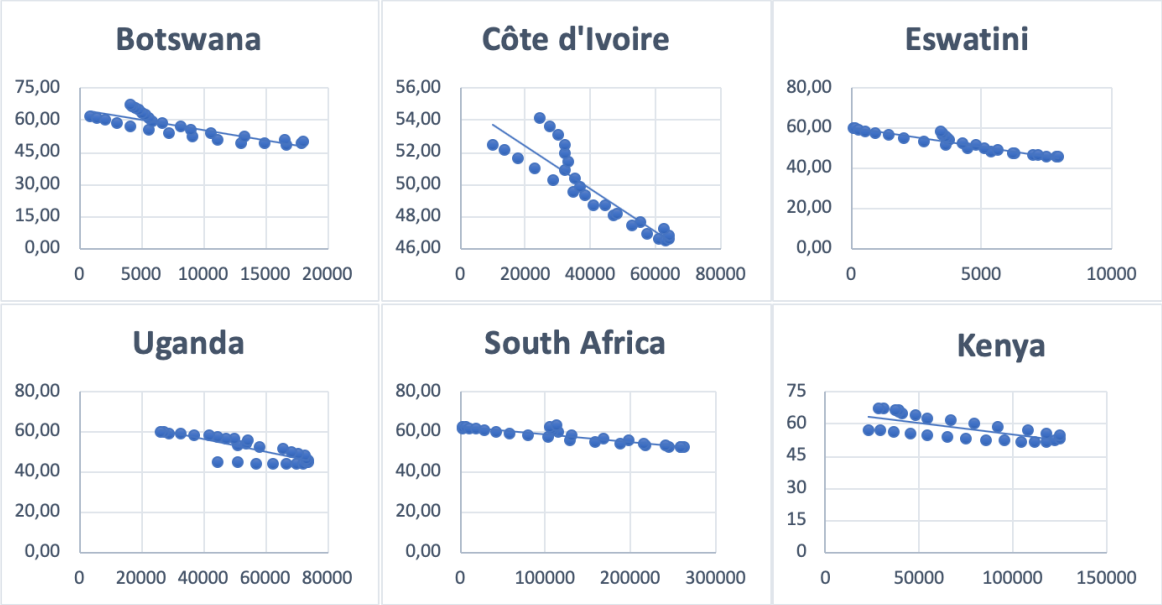

Figure 3: Life expectancy at birth (in number of years lived) on ARD (number of deaths).

2.2 AIDS as a natural experiment

In this thesis, the AIDS epidemic in Sub-Saharan Africa will be used as a natural experiment due to the exogenous effect we assume it has on life expectancy at birth. We analyse life expectancy during the most severe years of the AIDS epidemic and assume that AIDS is a main reason behind the decline in life expectancy at birth. The latter increase is assumed to be a consequence of AIDS-related deaths being reduced because of ART. This is an assumption made for this thesis, it does not ascertain that the effect is exogenous since many other factors could have affected life expectancy. Politically unstable countries are excluded from the sample (see section 4.2) to minimize the effect on life expectancy caused by for example conflicts. Still, a more country specific analysis would have to be made to be able to claim causality between ARD and life expectancy at birth and to assure an exogenous effect of AIDS on life expectancy.

The effect of AIDS is not a perfect natural experiment due to the possibility of endogeneity, we do however make the assumption of an exogenous effect of AIDS on life expectancy. The assumption is made with awareness of the limitations to the experiment and therefore also the accuracy of the results.

When analysing AIDS and economic growth, endogeneity and reversed causality cannot be ruled out. Even though AIDS is a disease which affects high-income countries as well as low-income countries, exogeneity cannot be assured in the regression analysis. In low developed countries, as this thesis studies, information about the spread and treatment is poor. In high-income countries on the other hand, medication and information is well-known and accessible, giving these countries a better basis to impede the AIDS epidemic. There is an endogeneity issue present since the ability to manage an epidemic is an important factor which affects the number of AIDS victims. The fact that countries in Sub-Saharan Africa struggle to access antiretroviral treatment due to insufficient funds is an important reason why reversed causality is also possible. This would mean that the dependent variable, economic growth, affects the independent variable, AIDS-related deaths. Due to this dilemma of reversed causality and endogeneity between AIDS and economic growth, causality cannot be established. Nonetheless, regressions using ARD as the variables of interest will be conducted to evaluate if it does have any significant effect on economic growth directly, and not only through the assumed effect on life expectancy. As stated, this will be analysed cautiously and with respect to the limitations that follow when exogeneity cannot be assured.

The variable which will be used when estimating the effect of AIDS is AIDS-related deaths (ARD). The rationale behind choosing ARD instead of looking at HIV prevalence mainly relates to the earlier mentioned introduction of antiretroviral treatment in the late 1990, which completely altered the standards of living for HIV-victims in two ways. First off, it stops the virus from developing to AIDS, consequently decreasing mortality rates. Simultaneous to this, it also decreases morbidity amongst the infected (WHO, 2019; U.S National Library of Medicine, 2016). The introduction of ART would naturally skew any regression results including HIV/AIDS-victims as a variable and concurrently make it difficult to extrapolate the effect with the dependent GDP-variable, hence the variable ARD is used instead. By doing so, the endogenous ART-effect is minimized and made more exogenous to the model, improving inference on the results.

3. Theoretical framework

3.1 Solow growth model

The textbook Solow growth model is centred around the production function, using only two key inputs to explain economic growth, capital and labour. The output, growth, increases with both the variables, and the function is seen to be subject to the law of diminishing marginal productivity. Below (equation 1) is the expression for the Solow model where the output growth *Y,* is a function *f* of the inputs*,* capital *K*, human capital *H,* productivity *A*, and labour L, where *AL* gives the productive labour force, as specified by Acemoglu (2011).

(1) $$Y = f(K, H, AL)$$

Human capital *H* represents health in our thesis, health is in turn captured by the measure for health, life expectancy at birth. The variable AIDS-related deaths is, through its assumed effect on life expectancy therefore affecting *H*. Education is also commonly included in human capital, but due to data limitations it is not accounted for. Not accounting for education is a deficiency with the thesis since AIDS possibly affected education as well. It is likely that for example, attendance from both students and teachers decreased during the worst years of the epidemic.

The Solow model includes technology/productivity *A* as a key variable for growth, a part not explained by changes in capital or labour. Productivity improvement, usually through exogenous technological breakthroughs, is the way to achieve long term growth since this is what drives *A,* according to the model. With constant productivity improvement, the growth rate will not return to previous steady state, as it would by changing labour or capital, instead it will shift the steady state upwards which generates long term growth. This thesis does not analyse the steady state or long term growth, technology will therefore be accounted for as exogenous in our model and *A* will be assumed as a constant.

The following Cobb-Douglas production function (equation 2) will be used to explain how growth has been affected by the changing life expectancy at birth in Sub-Saharan Africa, which is assumed to have been reduced by AIDS and then improved by the introduction of ART. When applying the production function in this thesis *Y* represents economic growth, *K* is the capital formation, *H* represents life expectancy or ARD, *A* is accounted for as exogenous and

L is the total labour force. We analyse the cross-country dimension *i*, and the cross-temporal dimension *t.* The applied model will be further presented in section 5.3.

(2) $$Y_{it} = K_{it}^{\alpha} H_{it}^{\beta} A L_{it}^{1-\alpha-\beta}$$

Acemoglu (2013) claimed that the Solow model is an efficient framework for proximate causes of growth, but that the model does not look at fundamental causes which underlie the proximate causes. This thesis aims to investigate if the changes in life expectancy at birth in Sub-Saharan Africa have had an impact on growth. Hence, the proximate causes for growth that the Solow growth model provides are adequate for the intention of this thesis.

3.2 Literature review

This section reviews earlier studies within our research area. Beginning with studies regarding the relationship between health and economic growth, followed by studies specifically on HIV/AIDS and economic growth.

Lorentzen, McMillian and Wacziarg (2008) investigated how adult mortality affects growth by examining the partial correlations between economic growth of income per capita with various measures for mortality. To analyse the cross-country data and the cross-sectional data for India, OLS regressions were conducted as well as an extended version which included a dummy variable for Sub-Saharan Africa.

Lorentzen, et al. (2005) studied three channels through which adult mortality affects economic growth; investment, human capital accumulation and fertility rate. The authors found that adult mortality causes people to change their behaviour in ways which disfavour growth. Investment rates and human capital accumulation decline while fertility increases, subsequently children increase in numbers yet quality, as measured by human capital accumulation, is on average lowered. All these three channels generally have a negative impact on economic growth. The authors claimed that adult mortality in Africa is a poverty trap and that in a statistical sense it can even account for the entire growth tragedy in Sub-Saharan Africa from 1960 to 2000.

Bloom and Canning (2008) explored the links between demographic change and economic growth. The econometric approach adopted was fundamentally the production function. Bloom

and Canning found that reductions in child- and infant mortality rate will lead to a reduction in fertility rate and subsequent reduce the dependency ratio. The dependency ratio is the ratio between the non-working-age cohort (0-14 years old and 65 years old and older) and the working-age cohort (15-64 years old) which defines the age structure of a region and is useful to estimate the burden the working-age cohort carries.

Given that the country is able to absorb the increased labour they can achieve demographic dividend from a low dependency ratio. The demographic dividend is a state of high labour-force participation and high savings rate, which has positive impacts on economic growth and will possibly encourage further investments in education (Bloom & Canning, 2008).

Acemoglu and Johnson (2007) estimated the impact health improvements has had on economic growth from 1940. The authors referred to earlier studies (Bloom & Sachs, 1998; Gallup & Sachs, 2001; WHO, 2001; Alleyne & Cohen, 2002; Bloom & Canning, 2005; Lorentzen, McMillan, & Wacziarg, 2005) which all share a consensus that improving health is more than a social objective, it also has a positive effect on economic growth. Acemoglu and Johnson on the other hand, claimed that the previous studies, showing a strong correlation between health measurements and economic growth, have not established a causal effect between diseases and GDP growth since other omitted factors have not been accounted for.

> Since countries suffering from short life expectancy and ill health are also disadvantaged in other ways (and often this is the reason for their poor health outcomes), such macro studies may be capturing the negative effects of these other, often omitted, disadvantages. (Acemoglu and Johnson, 2007, p. 926)

The authors reached the conclusion that increasing life expectancy cannot be proven to lead to faster economic growth of income per capita. They thereby also claimed that it is doubtful that health improvements have an important role in increasing economic growth.

In 1940, the major mortal diseases that Acemoglu and Johnson focused on, mainly affected children. The authors explicitly stated that the result of the study could be different if they had investigated a time of HIV prevalence. Since HIV/AIDS affects the working-age cohort as well, the virus would likely have a negative impact on economic growth (Acemoglu & Johnson, 2007).

In a study regarding HIV/AIDS and economic growth, Lovász and Schipp (2009) tested the hypothesis that HIV had a sustainable negative impact on economic growth rate per capita in Sub-Saharan Africa. The authors used health capital as a determinant of human capital, a variable postulated to be a consequence of material standard of living, health care expenditures and the incidence of contagious diseases. Lovász and Schipp reached the conclusion that HIV prevalence has had a significant negative impact on GDP per capita. Sub-Saharan Africa had on average an HIV prevalence rate of 5% and the economic growth rate was reduced by over two percentage points each year. Hence, in a scenario without HIV prevalence the region is likely to have had a growth rate more similar to South and East Asia. The authors therefore claimed that their hypothesis was true and that HIV/AIDS caused a negative impact on growth and that this negative effect has been sustainable in Sub-Saharan Africa.

Dixon, McDonald and Roberts (2001) tested the hypothesis if epidemics, such as AIDS, show any correlation with growth. In other words, if the effect is spuriously either positive or negative. Contrary to their hypothesis, the results from their econometric data analysis clearly demonstrate how the epidemic, during the latter so-called peak stage where the loss of lives start to impact social and economic interactions, has had a significant effect on economic development. Up until reaching that stage however the hypothesis was confirmed, as a clear correlation between HIV and economic development could not be established.

Alwyn Young (2005) simulated the effects AIDS will have on the future standards of living in Africa. The author referred to a Solow based model in his analysis regarding population growth, and behavioural models, to analyse how the behaviour of households change when facing a health crisis. Two competing effects are emphasized, fertility and human capital accumulation, between which Young found that fertility is the dominant effect on the future standards of living. The effect on fertility is negative, subsequently the population growth rate naturally declined which improves the dependency ratio. The author stated that the negative effect on population growth has endowed the region with extra resources which creates expectations on a future increase of consumption and therefore also of economic growth. Young concluded his article by clarifying that AIDS has been a humanitarian disaster, however not an economic one, seeing as the effect on economic growth was shown to be positive.

Among the studies presented, there is clearly no consensus regarding the effect HIV/AIDS, and health in general, has had on economic growth. The main mechanisms through which the

authors consider health influences economic growth are changes in productivity, behaviour and demography. On one hand, population is decreased and there are less people to share the resources with. On the other hand, the working-age cohort is reduced which decreases productivity. Which out of these effects is the dominant one in our model, will be estimated when reaching the result of how life expectancy, as a measure for health, affects economic growth.

4. Data

4.1 Data source

The data used in this study is mainly retrieved from the World Development Indicators (2019), assembled by the World Bank, the International Monetary Fund (IMF) (2019), and UNAIDS (2018). The cross-country data for the following variables have been collected from the World Development Indicator; life expectancy at birth, GDP (constant LCU), GDP (World growth %), total labour force, and gross fixed capital formation.

For the AIDS-related deaths the data was collected from UNAIDS (2018) and the commodity prices have been retrieved from the IMF (2019). Due to lack of data on commodity prices for Botswana and Eswatini from 1990 to 2000, these years have been approximated by using the same percentage change of a comparable country as a proxy. Togo is used as a proxy for Eswatini, as both these countries are highly dependent on cotton exports, and Namibia is used as a proxy for Botswana since both mainly export diamonds (Globalis, 2013-2018).

4.2 Sample countries and time frame

The observations span from 1990-2017, a wide span which has the advantage of clearly illustrating the entire course of the HIV/AIDS epidemic, starting from a low point and expanding until reaching a peak period. This period is then followed by a clear stalled decline in AIDS-related deaths after the introduction of antiretroviral treatment plans for HIV-victims (Figure 1). By including these transitions in the course of the epidemic, the explicit effect of the changes in life expectancy, through the effect of ARD, on economic growth is made applicable to our model.

According to the UN Development Program (2013) 46 African countries are considered Sub-Saharan, out of these the sample for this thesis contains the following 16 countries: Botswana, Burkina Faso, Cameroon, Côte d'Ivoire, Eswatini, Gabon, Kenya, Malawi, Mozambique, Namibia, Nigeria, Rwanda, South Africa, Tanzania, Togo & Uganda (table 1).

Due to shortcomings in reported data the following nations have been eliminated from the study: Angola, Equatorial Guinea, Eritrea, Ethiopia, Gambia, Guinea Bissau, Lesotho, Liberia, Sudan and Zambia. Neither have the smaller islands been analysed, such as Cape Verde, Comoros, Mauritius, Sao Tome and Principe, and Seychelles.

To isolate the effect ARD has had on life expectancy, and in turn, economic growth, Burundi, Central African Republic, Congo Democratic Republic, Somalia, South Sudan and Zimbabwe are not included due to e.g. civil wars and general political- and economic instability during the period of observation (Fragile State Index, n.d.).

Lastly, several countries were excluded due to low aggregated ARD during the studied period of time. The reason for this is that the scale of the effect on the dependent variable is otherwise limited, making it difficult to relate any change in ARD to changes in the economic growth. For a generalized epidemic, the HIV prevalence rate is typically over 5% in subpopulations whilst remaining below 1% in the general population (UNAIDS 2011). Low is in this thesis therefore defined, in line with the definition of a generalized epidemic, as an aggregated sum of 5% ARD relative to the total labour force in the respective country. Equation 3 shows that countries with an aggregated sum of ARD divided by labour above, or equal to, 0.05 for the total time period (1990-2017) are included, where t represents each individual year. This excludes Benin, Chad, Congo, Ghana, Guinea, Madagascar, Mali, Mauritania, Niger, Senegal and Sierra Leone from the sample (World Bank Indicators & UNAIDS). The countries which lie above the 5% limit, and which were neither excluded due to lack of data nor instability, are presented in table 1 with their respective aggregated percentages.

(3) $$0.05 \leq \Sigma_{i=t}^{T} \frac{ARD_t}{LABOUR_t} + \frac{ARD_{t+1}}{LABOUR_{t+1}} +, \dots, + \frac{ARD_T}{LABOUR_T}$$

Table 1: Percent of ARD/Labour between 1990-2017. *Refer to equation 3.

Country	**Labour force 2017 (total working age)**	**ARD/Labour in % (sum 1990-2017)***

Botswana	1131720	32,17
Burkina Faso	6995457	5,29
Cameroon	10499670	6,71
Côte d'Ivoire	8015967	17,56
Eswatini	466259	32,59
Gabon	667636	7,11
Kenya	19352405	15,66
Malawi	8029723	19,52
Mozambique	12898678	15,12
Namibia	989579	17,45
Nigeria	58958901	7,99
Rwanda	6296321	6,88
South Africa	22041176	19,75
Tanzania	26306084	11,30
Togo	3533538	5,15
Uganda	15839805	17,29

5. Methodology

5.1 Panel data

This thesis analyses cross-sectional and cross-temporal data with several observations per unit, hence a panel data analysis is conducted. The panel data covers the sample of 16 Sub-Saharan countries (table 1) over 28 years and includes several control variables in addition to the dependent variable and the variables of interest (section 5.2).

We apply Driscoll-Kraay standard errors, an alternative to robust standard errors which is used with fixed effects regression. This alternative standard error is consistent even if there is heteroscedasticity, cross-sectional dependence and/or autocorrelation, issues which are indicated to be present in our model after running several tests (more on this in section 6.2). A fixed effects model is applied to control for endogeneity in individual or time factors that could bias the outcome.

5.2 Variables

5.2.1 Dependent variable

The aim of this paper is to evaluate if life expectancy at birth, as a measure for health, has had any significant effect on economic growth in Sub-Saharan Africa. Hence, the dependent variable is in accordance with prior studies related to the Gross Domestic Product (GDP) of the panel countries.

This study utilizes GDP in constant local currency to measure the effect in real terms on GDP growth instead of in relative terms, as with the commonly used GDP per capita (GDP/population) measure (Kahende & Hoch, 2008; Bloom & Mahal, 1997; Dixon, McDonald & Roberts, 2001). By using constant local currency, the true growth of the series of observations is captured, i.e. the GDP is adjusted for the effects of price inflation by having a fixed base year.

As with most of the variables used in this study, the natural logarithm is most appropriate to account for differences in the magnitudes of GDP between the subjected countries. Lastly, it is important to note that the dependent variable is, similarly to the independent variables, denoted as a differential to capture growth (equation 4).

(4) $$\Delta lnY = lnY_t - lnY_{t-1}$$

When choosing the dependent variable as being measured by GDP in constant local currency, instead of GDP per capita, a deviation from what has been common practice in the research area was made (see e.g. Lovász & Schipp, 2009; Kahende & Hoch, 2008; Dixon, McDonald & Roberts, 2001; Bloom & Mahal, 1997; Ashraf, Weil & Wilde, 2013). The justification behind this is that the study intends to look at the direct effect of ARD on GDP growth in absolute terms instead of in relative terms. If ARD had been used in combination with GDP per capita, there would have been a direct effect on the numerator which would perhaps cancel out the effect on the denominator, consequently biasing the possible causality between these variables.

5.2.2 Variable of interest

As with previous studies in the area, a variable controlling for the health level of each nation is highly relevant as the hypothesis of this study relies on the assumption that changes in the wellbeing of people has a significant effect on the dependent variable. Different measures to account for health have been used before, such as human capital and health expenditures, which have the advantage of being quite comprehensive in their grasp, yet are difficult to obtain reliable data for and are in ways subjective (Dixon, McDonald, & Roberts, 2001; Kahende & Hoch, 2008). In accordance with Acemoglu and Johnson (2008), the measure used to account for health development is life expectancy at birth.

Since the measure for health, life expectancy at birth, places more weight to recent years it is not a perfect measure for health in general, as it relies on the assumption that recent years better explain the present (WHO, 2006). This would in the case of extreme externalities, such as the outbreak of an epidemic or a civil war, skew the life expectancy at birth quite drastically and would thus be regarded as a weakness in the measure. For the purpose of this particular study however, this is rather seen as a strength as it highlights the effect of AIDS on GDP growth. Another positive aspect of the measure is that it is easily measured while also being relatively objective, making it one of the most frequently used measures.

Although prior studies within the covered research area have commonly regarded the total number of HIV/AIDS-victims as their variable of interest, this paper focus solely on the number of AIDS-related deaths (Kahende & Hoch, 2008; Dixon, McDonald & Roberts, 2001; Bloom & Mahal, 1997). AIDS-related deaths affect mortality directly and is therefore more relevant since the health measure used in this study is, as aforementioned, life expectancy at birth (see section 2.2 for further discussion). The variable is adjusted to the same format as the dependent variable, expressed in equation 5 below.

$$\Delta lnX = lnX_t - lnX_{t-1} \tag{5}$$

5.2.3 Solow input variables

Since the model used is based on the theoretical framework of the Solow Growth Model, discussed in section 3.1, variables accounting for changes in capital and labour naturally had to be included. The aim was thus to use the capital stock for each nation starting at year 1990 and add the capital investments for each year that followed, but because of data limitations this

measure could not be applied. Instead, a proxy for the capital stock, which still holds in the production function according to Solow's theory, was used. Similar to Lovász and Schipp (2009), the capital investments in our regression models are operationalized by gross fixed capital formation (denoted as FCF). This variable describes the net capital investments in the capital stock in the production function for the Solow growth model.

Next, data for total labour force was gathered for all 16 countries and 28 years, a straightforward measure denoted in absolute terms. Just as with the previous variables, these variables were adjusted as shown in equation 5.

An important issue to be aware of is the endogeneity that capital investments could have in the model as it should display a $\beta \neq 0$ and $\text{cov}(CapF, Labour) \neq 0$, which would suggest that the estimated marginal effect is not equal the true marginal effect β. The same applies for labour given the same conditions.

5.2.4 Control variables

In order to minimize omitted variable bias, additional relevant control variables were included. Seeing as the economies of the countries included in this thesis are dependent on commodity exports, a variable accounting for commodity price fluctuations was relevant to include. The commodity net export price index from the IMF was therefore used. This measure accounts for the different weights each commodity has in the specific country's export and correlates this with the change in the international price of these commodities. Since the data is indexed to year 2012 a reindexing had to be conducted to the start year of this study, 1990, to more persistently show the effects of these price fluctuations. As data for two countries was missing for parts of the studied timeframe, these were approximated by looking at countries with comparable commodity export composition and applying an averaged index change to the countries with missing observations (section 4.1).

Another control variable that was relevant to include was World GDP growth. By including this variable, any global economic growth fluctuations were made endogenous to the model.

Table 2: Variables and respective descriptions.

Variable	Label	Description	Unit	Year

Dependent variable				
lndiffGDP	GDP in constant local currency	Growth of GDP in constant local currency	Natural log(%)	1991-2017
Independent variables				
LifeExp	Life Expectancy	Change in life expectancy at birth	Natural log(%)	1991-2017
ARD	Aids-related deaths	Change in Aids-related deaths	Natural log(%)	1991-2017
CapF	Capital Formation	Change in Gross Fixed Capital Formation	Natural log(%)	1991-2017
CapFLag	Capital Formation, lagged	See section 5.2.3.	Natural log(%)	1992-2017
Lab	Labour	Change in labour for working age population	Natural log(%)	1991-2017
LabLag	Labour, lagged	See section 5.2.3.	Natural log(%)	1992-2017
WorldGDP	GDP growth World	GDP growth for the World	Natural log(%)	1991-2017
CPC	Commodity Price	Change in Commodity Price	%, Index (Base 1990)	1991-2017
CPCLag	Commodity Price, lagged	See section 5.2.4.	%, Index (Base 1990)	1992-2017

5.3 Econometric model

A regression analysis has been conducted for our econometric model, one basic followed by 3 with modifications. The short model (equation 9) is constructed from the Solow growth model, and derived from the Cobb-Douglas production function as seen below in equation 6-8. The short model aims to evaluate how the economic growth has been affected by the proximate causes that the Solow model provides; capital, human capital, which includes life expectancy at birth or ARD, productive labour, and the error term ε. For a detailed description of the notation of the variables in the following regressions see table 2.

Here follows a derivation of the Cobb-Douglas production function leading to the model constructed for the purpose of this study, where *i* represents the cross-sectional dimension and

t the cross-temporal dimension. *K* stands for capital formation and *H* represents human capital which is captured mainly by health, ARD or life expectancy. *AL* is the productive labour force, where *A* is considered a constant in the model.

(6) $$Y_{it} = K_{it}^{\alpha} H_{it}^{\beta} AL_{it}^{1-\alpha-\beta}$$

(7) $$lnY_{it} = \alpha\ ln\ (K_{it})\ + \beta ln\ (H_{it}) + (1 - \alpha - \beta) ln\ (AL_{it})$$

(8) $$\Delta lnY_{it} = \Delta\alpha ln(K_{it}) + \Delta\beta ln(H_{it}) + \Delta(1 - \alpha - \beta) ln(AL_{it})$$

As mentioned in section 3.1, life expectancy is assumed to capture health which is included in human capital *H.* ARD is in turn assumed to have a direct effect on *H* through its effect on life expectancy. Life expectancy at birth and AIDS-related deaths are both variables of interest in the study, and are respectively substituted with $\Delta\beta ln(H_{it})$in the econometric model conducted for this study (equation 9-12). As mentioned, productivity *A* is accounted for as exogenous, a constant, and therefore not included in the applied model. An argument can be made that AIDS affects *A* as well, but since the focus for this thesis is health, we consider *A* to be a constant and only study the effect through *H.* To clarify that *H* is the variable of interest this term will be after the intercept coefficient in the regression in the following equations. The error term denoted by ε represents the unexplained effect on the dependent variable, that is, all that is not captured by the independent variables.

(9) $$\Delta lnGDP_t = \beta_0 + \beta_1 \Delta ln(H_t) + \beta_2 \Delta ln(CapF_t) + \beta_3 \Delta ln(Lab_t) + \varepsilon_t$$

The calculation for the separate independent variables in the equation is explained by equation 5. The two variables of interest will not be included in the same regression since they are expected to be correlated, and are both measures of health. We will present two sets of regressions for the model, one using life expectancy at birth and one using ARD, in order to see what effect each variable separately has on the dependent variable GDP growth.

A modification to the short model is conducted by including the variables for capital and labour with a lag (equations 10-11). The lagged variables are included due to concerns regarding endogeneity in the model. It is important to note that although this helps to detect endogeneity in the model, it does not rule it out in any way.

(10) $$\Delta lnGDP_t = \beta_0 + \beta_1 \Delta ln(H_t) + \beta_2 \Delta ln(CapFLag_{t-1}) + \beta_3 \Delta ln(LabLag_{t-1}) + \varepsilon_t$$

(11) $$\Delta lnGDP_t = \beta_0 + \beta_1 \Delta ln(H_t) + \beta_2 \Delta ln(CapF_t) + \beta_3 \Delta \ln(Lab_t) + \beta_4 \Delta ln(CapFLag_{t-1}) + \beta_5 \Delta ln(LabLag_{t-1}) + \varepsilon_t$$

To isolate the effect AIDS-related deaths and life expectancy at birth respectively have on economic growth, the long model which is an extension of the short model (equation 12), includes two additional control variables. The long model controls for World GDP growth and commodity prices, which both are expected to affect economic growth in the sampled Sub-Saharan countries.

(12) $$\Delta lnGDP_t = \beta_0 + \beta_1 \Delta ln(H_t) + \beta_2 \Delta ln(CapF_t) + \beta_3 \Delta ln(Lab_t) + \beta_4 \Delta \ln(CapFLag_{t-1}) + \Delta ln(LabLag_{t-1}) + \beta_6 \Delta ln(WorldGDP_t) + \beta_7 \Delta lnCPC_t + \beta_8 \Delta lnCPCLag_{t-1} + \varepsilon$$

6. Results

6.1 Regression results

In this section, we will present the results from our econometric model, with all its specifications. The purpose of these 8 regressions is to see if the results hold to explain how life expectancy affects economic growth in the 16 selected Sub-Saharan countries. The number of observations vary from 432 to 416 due to lagged variables in model 2-4 and 6-8 (seen in tables 3 and 4). We accept a significance level of 10%.

The two variables of interest used, life expectancy at birth and AIDS-related deaths, are estimated and illustrated in individual regression tables (table 3 and 4, seen below). Each of these use the same control variables, specifications and both have GDP in constant local currency as the dependent variable. The short model is labelled as model 1 and 5 respectively, with modifications for lag in model 2-3 and 6-7. The long model of the regressions is labelled model 4 and 8 respectively. The results are explained below and presented in table 3 and 4. Firstly, we explain the results in the regression table 3 where life expectancy at birth is used as the variable of interest, then follows the results for the regressions with AIDS-related deaths as the variable of interest, seen in table 4.

Firstly, in model 1, represented by equation 9, life expectancy at birth is significant at a 10% significance level. The coefficient for life expectancy at birth β_1 is positive and indicates that

generally in the sample countries, a 1% increase in life expectancy at birth is associated with an increase in GDP growth of 0.245%. The inputs capital and labour are both positive, labour has a quite high coefficient but is not significant, whereas capital is significant at a 5% significance level and its coefficient is at a level of 0.117.

When using lagged control variables, we capture the dynamic in the model and take the fact that previous years possibly affect the next year into consideration. In these specifications, model 2 and 3, life expectancy at birth is significant even at a 5% level. Capital and labour are not statistically significant when they are lagged which indicates that the previous year does not have a significant effect on the GDP growth of the following year. In model 3 (equation 11) where capital and labour are included both in lagged form and not, capital displays a lower p-value. The adjusted R-squared is reduced to 0.064 in model 2 when including lagged variables. Hence, this specification to the short model does not explain the dependent variable in the same extension as model 1, without lagged variables, does. Model 3 on the other hand increases the R-squared value, providing a better explanation to the outcome.

Finally, the long model (model 4, equation 12) shows a significant effect on life expectancy at birth at a 5% significance level. The coefficient β_1 is interpreted as increasing the outcome variable GDP growth with 0.289% when life expectancy at birth increases with 1%. As in previous regressions for the short model, labour is not significant and capital is significant at a 5% significance level.

World growth is controlled for in the long model and is significant at a 1% significance level, its coefficient β_6 indicates an increase of 0.628% in GDP growth for the sample countries when the world growth increases with 1%. Commodity prices is included in the long model, both lagged and not, neither is significant at the accepted rejection level for this thesis. The coefficient for commodity prices are estimated to be negative when not lagged, and positive when lagged.

In the long model, we find the highest R-squared value, comparing the model specifications with life expectancy at birth as the variable of interest. This indicates that the inputs in the long model are important to add to better explain the outcome, economic growth.

Table 3: Models with life expectancy at birth as variable of interest, regressed on GDP (Constant Local Currency)

Regression table with GDP Constant Local Currency as dependent

	Model 1	Model 2	Model 3	Model 4
Life Expectancy	0.245*	0.403**	0.296**	0.289**
	(0.098)	(0.015)	(0.013)	(0.017)
Capital	0.117**		0.122**	0.120**
	(0.025)		(0.023)	(0.025)
Labor	0.767		0.299	0.245
	(0.119)		(0.195)	(0.273)
Capital, lagged		0.0144	0.0251	0.0258
		(0.602)	(0.189)	(0.202)
Labor, lagged		1.039	0.663	0.725
		(0.195)	(0.185)	(0.154)
GDP Growth World				0.628***
				(0.005)
Commodity Price				-0.108
				(0.130)
Commodity Price, l~d				0.0911
				(0.132)
Constant	0.0138	0.00974	0.00668	-0.0114
	(0.412)	(0.685)	(0.727)	(0.591)
Observations	432	416	416	416
Adjusted R-squared	0.1912	0.0640	0.2101	0.2356

p-values in parentheses
* p<0.10, ** p<0.05, *** p<0.01

When including ARD as the variable of interest (table 4), the short model specification (model 5, equation 9), shows that the variable of interest is insignificant at the 10% level, although with a p-value of 0.103. A negative coefficient is estimated in line with what a direct effect of increased mortality should result in. This would indicate that an increase in ARD would lead to a decrease in GDP growth. When lagged, ARD was still estimated as insignificant as shown in table A2 in the appendix. Both capital and labour in model 5 are significant at a 5% significance level and the coefficients indicate a positive effect on the dependent variable.

Labour displays a relatively strong coefficient of 0.9, while the coefficient on capital is weaker at 0.114.

Moving on to the second specification, namely model 6 (equation 10), where labour and capital are included in a lagged form. The p-value of the variable of interest, ARD, increases slightly in this specification, compared to model 5. In addition, the results show that the lagged variables are non-significant at any relevant rejection level. This means that the lagged variables do not have any statistically significant impact on the non-lagged dependent variable, namely GDP growth. Lastly, the R-squared value of 0.072 is lower when running this particular specification, relative to the other specifications, indicating that the inputs are less informative about the outcome.

In model specification 7 (equation 11), the p-value of ARD increases yet again to 0.159, but the coefficient stays negative at approximately the same level as earlier specifications. Both the non-lagged- and lagged variables for capital and labour are included here. Just as in model 5, the variables are significant at a 5%-level while the lagged variables are non-significant as in model 6. Nonetheless, it is interesting to discern the fact that the coefficients as well the p-values of most of the independent variables change quite substantially as compared to previous specifications. The signs in front of the coefficients are however kept intact.

To conclude, model 8 (equation 12) is the long specification of the model which includes an additional 3 control variables. Here, the p-value of the variable of interest drops to 0.104, and is thus, just as in model 5 randomly deviated from the 10% significance level. In comparison to model 7, the main difference relates to the added variables, although the labour variable does become less significant reaching only a 10%-significance level. The World GDP growth variable is estimated to be significant and has a positive coefficient. The commodity price variable estimates a negative coefficient, but is insignificant at the 10% level. A final remark is that the R-squared value for model 8 is the highest amongst all (model 1-8), signalling that its inputs explain the outcome, economic growth, better than the other specifications.

Important to note is that the p-value for the variable of interest ARD lies just above the 10% significance level in all four specifications and has similar coefficient estimates. Being this close to the chosen critical value level, no conclusions can be drawn as this deviation should

be seen as random, small changes or improvements in data would likely make the variable significant at the 10% level.

Table 4 - Models with ARD as variable of interest, regressed on GDP (Constant Local Currency)

Regression table with GDP Constant Local Currency as dependent

	Model 5	Model 6	Model 7	Model 8
AIDS-related deaths	-0.0373	-0.0514	-0.0418	-0.0397
	(0.103)	(0.128)	(0.159)	(0.104)
Capital	0.114**		0.121**	0.118**
	(0.023)		(0.020)	(0.022)
Labor	0.900**		0.450**	0.395*
	(0.040)		(0.040)	(0.071)
Capital, lagged		0.0109	0.0221	0.0228
		(0.701)	(0.219)	(0.238)
Labor, lagged		1.105	0.621	0.681
		(0.176)	(0.228)	(0.192)
GDP Growth World				0.614***
				(0.002)
Commodity Price				-0.111
				(0.113)
Commodity Price, l~d				0.0837
				(0.185)
Constant	0.0140	0.0134	0.00796	-0.00971
	(0.347)	(0.565)	(0.659)	(0.616)
Observations	432	416	416	416
Adjusted R-squared	0.1984	0.0716	0.2172	0.2415

p-values in parentheses
* p<0.10, ** p<0.05, *** p<0.01

6.2 Sensitivity analysis

To choose specification of the model, a Hausman test was conducted using the original short model specification A (appendix, table A1). The null hypothesis is that the individual-level effects of the estimators are adequately modelled by a random-effects model. Results show that the null hypothesis is clearly rejected at the 5% significance level, meaning that a fixed-effects

model is more appropriate to apply in this case. Random effects model has stronger demands on exogeneity and is therefore not appropriate for macro-level models, as the one in this study. The fixed effects model captures unobservable variables which change over time. This is of high importance since the countries included are heterogeneous in the aspect of for example institutions, social policies and weather, all factors which could influence the economic growth in each country differently.

Heterogeneity between ARD and labour is to be expected as both $\beta \neq 0$ and $cov(ARD, Labour) \neq 0$, an increase in ARD would naturally lead to a decrease in labour. Hence, multicollinearity is possibly present when including ARD together with labour. Multicollinearity can lead to skewed results since it may create a wider confidence interval. Subsequent, type 2 errors are likely to occur, resulting in failure to reject a false null hypothesis. However, multicollinearity does not interfere with the predictive power of the outcome in the model as a whole, but solely with the coefficient to the variables which have an interacting effect between them, which ARD and labour are expected to have in this study. This is possibly a reason why ARD is only close to significant in the regression when using it as the variable of interest. The insignificance can also be a consequence of insufficient data or low aggregated ARD in certain countries (section 4.2). ARD would possibly be significant had we only included countries with higher aggregated ARD than the used limit of 5%.

Omitted variable bias is expected in the short model specifications, model 1-3 and 5-7. This is the model in its simplest form and omitted variable bias is assumed to be present as relevant variables are excluded. In the long model, model 4 and 8, the higher R-squared values provide evidence that the additional control variables, world growth and commodity prices, are important inputs to explain the outcome. By including additional variables, the risk of omitted variable bias can be reduced.

In order to check for differences in the statistical dispersion of the variables included, a test for heteroscedasticity was conducted. The modified Wald test for groupwise heteroscedasticity in fixed effect regression models was the most appropriate and it clearly showed that a null hypothesis of homoscedasticity had to be rejected. Although not surprising, seeing as there are many potential exogenous country-specific factors which could alter variance, it is something that had to be dealt with since failing to do so could invalidate any statistical inference from our model.

Furthermore, a Breusch-Pagan LM test of cross-sectional dependence was conducted to control for any interdependencies between the cross-sectional units estimated in the panel data regressions (see appendix tables A3 and A5). A test best suited when T>N, meaning that the time dimension is larger than the cross-sectional dimension (De Hoyos & Sarafidis, 2006), which is the case in this study. The importance of this is emphasized by Hoechle (2007), who argues that "erroneously ignoring cross-sectional correlation in the estimation of panel models can lead to severely biased statistical results.", and states that this has been a commonly ignored aspect amongst practitioners. De Hoyos and Sarafidis (2006) also discuss this, and point to the fact that cross-sectional dependence is likely to be exhibited in the errors of panel data models due to common shocks being present together with unobserved components, which ultimately become part of the error term.

Results from the Breusch-Pagan LM test show that there is in fact cross-sectional dependence as the null hypothesis of there being no cross-sectional dependence is rejected at all significance levels. Finally, a modified Wald test for groupwise heteroscedasticity in fixed effect regression models (see appendix tables A4 and A6) was conducted to check for potential presence of heteroscedasticity in the variance of the variables. The null hypothesis was again rejected in this test, stating that there is an absence of homoscedasticity. In order to minimize the bias on the standard error that both the heteroscedasticity and contemporaneous correlation cause, Hoechle (2007) suggests that Driscoll & Kraay standard errors should be used. Developed by Driscoll & Kraay (1998) as a "nonparametric covariance matrix estimator which produces heteroscedasticity consistent standard errors that are robust to very general forms of spatial and temporal dependence.". As cited, it does not only control for the issues proven to be present, but also autocorrelation, which is assumed to be present to some extent in our model. Tables A9 and A10 test for autocorrelation in the two independent variables, and results show that there is a clear degree of serial correlation present at least for ARD.

7. Discussion

When health declines and affects adults, it will naturally have a direct negative effect on economic growth through the loss of productivity due to illness and/or death. This thesis concentrates on the health measure life expectancy at birth which does not account for how subjectively healthy the population was during their lifetime. Omitted variable bias may be

present since we do not account for how the time of illness affected growth. To adjust the model for this endogeneity problem, a more detailed analysis on health expenditures, productivity loss and welfare system would be required, as well as an analysis of other health measures. It is possible that the results would have been slightly different if another measure for health was applied, this would however entail making more subjective assumptions in constructing the measure (e.g. human capital).

AIDS takes lives from all age cohorts, which Acemoglu and Johnson (2005) claimed was a possible reason why the AIDS epidemic would impact growth negatively, in contrast to other mortal diseases which mainly affect children and elderly. The dependency ratio is one of the mechanism that can be used to confirm that AIDS affects growth negatively due to its increase in adult mortality. A reduction in the working-age cohort leaves the society unbalanced with a redundancy of children and elderly, giving the productive adults more people to provide for. In such a scenario, economic growth is expected to decrease, as the results in this thesis with ARD as the variable of interest shows (table 4). ARD is close to significant at our acceptable rejection level 10%, with a p-value of 0.104 in the long model, and it can therefore at least be assumed that the estimated coefficient has the correct sign, negative. With only 0.004% from significance, it is possible that with additional and/or improved data, ARD would have been significant.

Young (2005) estimated that the dependency ratio would decrease due to less people in all age cohorts when analysing the AIDS epidemic. However, the positive effect that Young estimated this would have on economic growth is not supported by this thesis since the results show a negative effect on economic growth. Hence, according to the results in this study, the effect from the reduction in productive adults is stronger than the effect of the reduction of population overall, which would result in extra resources relative to population. The prospects of improved growth that Young presented is however possible, but first when the epidemic settles. The high fertility, estimated by Lorentzen, et al. (2008), and sudden fall in mortality when ART is introduced will create a large future working-age cohort in comparison to elderly. In such a scenario, the dependency ratio would fall as Young (2005) estimated. Bloom and Canning (2008) discussed that in a situation with a decreasing dependency ratio, the region may achieve demographic dividend. To properly confirm this statement a more detailed analysis of the demographic change post the epidemic would have to be conducted.

The almost significant estimated coefficient for AIDS-related deaths (section 6.1) indicate a negative effect on growth, which can also be a consequence of the change in behaviour that adult mortality is associated with, in accordance with Lorentzen, et al. (2008). Death at the peak of the productive years is associated with riskier behaviour, higher fertility and lower investment rates, both for physical capital and human capital accumulation (Lorentzen, et al. 2008). Young (2005) also discussed how the behaviour of people changed during the worst years of the HIV/AIDS epidemic. In contrast though, the author found that there was a reduction in fertility which would have a positive impact on economic growth due to a redundancy of resources per capita, this is as stated, not supported by the results in this study. Even though ARD alone was only almost significant, we do expect it to be correlated with life expectancy which in turn has a statistically significant effect on economic growth. Hence, AIDS is assumed to have reduced life expectancy at birth and in extension therefore also decreased economic growth in Sub-Saharan Africa.

AIDS does not only affect productive adults though, children are infected by the virus as well. Children are one of the important sources to major reductions in life expectancy at birth. Death in early ages has a large impact on this particular health measure, since it is a calculation of the average years a new-born is expected to live (OECD 2017). The loss of children/infants does not cause a direct loss in productivity. Notwithstanding the fact that it does affect the future standards of living which Young (2005) discussed.

Life expectancy at birth is expected to have positive coefficients, as it does in our model, associating it with a positive effect on GDP growth. Life expectancy at birth is, as mentioned, significant in all the models, but since the R-squared value is highest in the long model (model 4) it is the one considered to explain the outcome the most. A decrease of 1% in life expectancy at birth is associated with a decline in GDP growth with 0.289% according to model 4.

In Botswana, the decrease of 21% in life expectancy at birth during the worst years of the epidemic, 1990-2001, mentioned in section 2.1, would then imply a decline in GDP growth of 6.07%. The followed increase in life expectancy from 2001-2017, partially expected to be due to ART, would according to our results have increased GDP growth with 10.6%. We cannot conclude that AIDS bore the full responsibility for the decline, though it is expected to be partly liable. Since ARD is not significant in our model, its estimated coefficient is not analysed further than having a negative impact and the impact it has through life expectancy at birth.

Lovász and Schipp (2009) estimated a 2% decline in GDP growth per year associated with the HIV/AIDS epidemic in Sub-Saharan Africa. Lovász and Schipp (2009), Over (1992) and Lorentzen et al. (2005) claimed that HIV/AIDS is accountable for the growth tragedy in Africa (section 3.3). The authors stated that in a scenario without HIV/AIDS, the economic growth in the region would be more similar to South and East Asia. This statement was validated by the model applied in this thesis, as it showed that ARD did explain a major part of the decrease in life expectancy in the sample countries, which subsequently had a negative effect on economic growth.

Acemoglu and Johnson (2007) discussed that omitted disadvantages in low developed countries, which affect economic growth and health, are present, and therefore causality cannot be established. HIV/AIDS affects high-income countries as well, but not to the same extent as low-income countries. In low-income countries as this study covers, the lack of information about the spread of HIV/AIDS and the lack of funds to provide treatment to the population are two major disadvantages which cause the epidemic to be more severe. Anyone could become infected, though the probability is higher in developing countries and the treatment is not as accessible, subsequent, there will be more AIDS-related deaths. Reversed causality and/or omitted disadvantages can therefore not be excluded when using AIDS as a natural experiment. The omitted disadvantages which are discussed by Acemoglu and Johnson (2007) causes possible reversed causality and the results of this thesis must therefore be considered with cautious.

Considering that the world is continuously becoming more globalised and interdependent on multilateral trade, price fluctuations were hypothesized to have a measurable effect on GDP growth for the selected Sub-Saharan countries. Surprisingly however, estimates retrieved on this variable were not significant even on a 10% significance level, yet had as predicted a negative coefficient.

The secondary data sources, from where the data for this thesis has been collected, World Bank, IMF and UNAIDS, are all well-established and recognized amongst prior studies, they should therefore be viewed as reliable. However, in low-income countries, as those this report studies, the primary data which the countries report to the secondary sources is not assured to be of

high quality. This is due to the risk of being affected by the quality of the government accounting systems and the level of corruption in the country (World Bank, 2019).

The data on AIDS-related deaths are also in risk of being faulty due to incorrect reports about causes of deaths or even the lack of any reports on deaths. In this aspect, the thesis faces limitations which compromises the truthfulness of the result.

8. Conclusion

The main purpose of this thesis was to analyse if life expectancy at birth has an effect on economic growth. In a separate model specification, the assumed exogenous effect that AIDS has had on life expectancy in Sub-Saharan Africa is used as a natural experiment and its effect on economic growth is estimated.

As the statistical significance of AIDS-related deaths randomly deviated from the chosen 10%-significance threshold, one careful conclusion that can be drawn is that AIDS did indeed cause a decrease in GDP growth in the sampled countries. A significant effect that we on the other hand can establish, is that the decline in life expectancy from 1990 until early 21st century decreased economic growth in Sub-Saharan Africa. We do not however claim causality due to endogeneity issues, as discussed in section 2.2. Yet given the assumptions made in this thesis, the results do show a relationship between life expectancy and economic growth. When comparing the model specifications, the importance of the control variables role to explain the outcome is made visible. The long model has the highest R-squared value and world growth is, as expected, an essential and statistically significant input to explain the outcome.

The discussed reasons behind the findings include productivity loss, demographic change and change in behaviour, mainly in a disfavouring way of economic growth, these are supported by previous work within the research area. When life expectancy at birth declines due to ARD, people tend to invest riskier, have more children and human capital accumulation declines, which is generally negatively associated with economic growth. Having obtained these findings, we can hereby respond to the research question in the introduction with the following. Yes, according to the results of this thesis, and given its assumptions, life expectancy does affect economic growth in Sub-Saharan Africa. In consent with Over (1992) and Lovász and Schipp (2009), it is very likely that Sub-Saharan Africa would have experienced another growth

development if the region had not been a victim of health issues which decrease life expectancy, such as the AIDS epidemic.

One of the main strengths of this study was the time-dimension which, in comparison with previous research in the area, was able to capture both the increase, peak, and subsequent decline of ARD in the Sub-Saharan region. In addition to this, the study also differentiates itself by applying a more direct measure to capture the impact of the HIV/AIDS-epidemic, by looking at the number of deceased AIDS victims instead of a more general HIV prevalence measure.

Despite these strengths, it is important to note that several areas can be improved upon and complemented in further research. One obvious way to develop the study would be to extend the cross-sectional dimension by including more countries. Another way could be to look at other pandemics which have severely affected productivity in the afflicted countries through increased mortality and decreased life expectancy. By doing so, a more generalizable conclusion may be reached. Lastly, future research could also complement this thesis by looking at the post-epidemic long-term effects of life expectancy on economic growth, by extending the time-dimension.

9. References

Acemoglu, D. (2013). Economic growth and development in the undergraduate curriculum. *The Journal of Economic Education*, *44*(2), pp.169-177.

Acemoglu, D., & Johnson, S. (2007). Disease and development: the effect of life expectancy on economic growth. *Journal of political Economy*, *115*(6), 925-985.

Acemoglu, D. (2011). *Economic Growth: Lecture 4, The Solow Growth Model and the Data*. [PowerPoint-presentation]. Retrieved 2019.05.20 from MIT https://economics.mit.edu/files/7183

Ashraf, Q. H., Weil, D. N., & Wilde, J. (2013). The effect of fertility reduction on economic growth. *Population and development review*, *39*(1), 97-130.

Avert. (2018). History of HIV and AIDS overview. Retrieved from https://www.avert.org/professionals/history-hiv-aids/overview

Avert. (2017). HIV and AIDS in East and Southern Africa regional overview. Retrieved from https://www.avert.org/professionals/hiv-around-world/sub-saharan-africa

Avert. (n.d). HIV Timeline. Retrieved from https://timeline.avert.org/

Bloom, D. E., & Mahal, A. S. (1997). Does the AIDS epidemic threaten economic growth?. *Journal of Econometrics*, *77*(1), 105-124.

Bloom, D. E., & Canning, D. (2008). *Global demographic change: Dimensions and economic significance* (No. w10817). National Bureau of Economic Research.

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *The stata journal*, *6*(4), 482-496.

De Janvry, A., & Sadoulet, E. (2015). *Development economics: Theory and practice*. Routledge.

Dixon, S., McDonald, S., & Roberts, J. (2001). AIDS and economic growth in Africa: a panel data analysis. *Journal of International Development: The Journal of the Development Studies Association*, *13*(4), 411-426.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of economics and statistics*, *80*(4), 549-560.

Fragile State Index. (n.d.). Global Data. Retrieved 2019.04.20 from https://fragilestatesindex.org/

Globalis. (2016). *Swaziland*. Retrieved 2019.04.30 from https://www.globalis.se/Laender/Swaziland.

Globalis. (2014). *Botswana*. Retrieved 2019.04.30 from https://www.globalis.se/Laender/Botswana

Globalis. (2018). *Togo*. Retrieved 2019.04.30 from https://www.globalis.se/Laender/Togo

Globalis. (2013). *Namibia*. Retrieved 2019.04.30 from https://www.globalis.se/Laender/Namibia

Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal*, *7*(3), 281-312.

Kahende, J., & Hoch, I. (2008). HIV/AIDS and economic development: evidence from thirty-nine sub-Saharan countries. *Perspectives on Global Development and Technology*, *7*(2), 151-173.

Lorentzen, P., McMillan, J., & Wacziarg, R. (2008). Death and development. *Journal of economic growth*, *13*(2), 81-124.

Lovász, E., & Schipp, B. (2009). THE IMPACT OF HIV/AIDS ON ECONOMIC GROWTH IN SUB-SAHARAN AFRICA 1. *South African journal of economics*, *77*(2), 245-256.

OECD. (2017). CO1.2: Life expectancy at birth. Retrieved 2019.05.17 from https://www.oecd.org/els/family/CO_1_2_Life_expectancy_at_birth.pdf

Over, M. (1992). The macroeconomic impact of AIDS in sub-Saharan Africa.

UNAIDS, U. (2011). UNAIDS terminology guideline 2011. Retrieved 2019.05.16

U.S National Library of Medicine. (2016). Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV. Retrieved 2019.05.09 from https://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv/2/introduction

World Bank. (2019). Gross capital formation. Retrieved 19.04.20 from https://data.worldbank.org/indicator/NE.GDI.TOTL.CD?locations=ZG

World Health Organization. (2006). WHO Stat 2006 Definitions and Metadata. Retrieved 19.05.20 from https://www.who.int/whosis/whostat2006DefinitionsAndMetadata.pdf

World Health Organization. (2016). Top 10 causes of death. Retrieved 2019.05.08 from https://www.who.int/gho/mortality_burden_disease/causes_death/top_10/en/

World Health Organization. (n.d.). Health information and knowledge management. Retrieved from https://www.afro.who.int/health-topics/health-information-and-knowledge-management

World Health Organization. (2019). Global Health Observatory (GHO) data. Antiretroviral therapy (ART) coverage among all age groups. Retrieved 2019.05.09 from https://www.who.int/gho/hiv/epidemic_response/ART_text/en/

Young, A. (2005). The gift of the dying: The tragedy of AIDS and the welfare of future African generations. *The Quarterly Journal of Economics*, *120*(2), 423-466.

10. Appendix

Table A1: Hausman test

	Coefficients (b) fixed_eff	(B) random_eff	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
lLifeexpDiff	**.2454826**	**.3404663**	**-.0949837**	**.050528**
lFCFdiff	**.1165521**	**.1206622**	**-.0041101**	**.0010857**
lLabordiff	**.766852**	**.6928984**	**.0739536**	**.0841118**

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(**3**) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= **15.41**
Prob>chi2 = **0.0015**

Table A2: Effect on Life expectancy from ARD

lLifeexpDiff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lARDdiff	**-.0500634**	**.0044829**	**-11.17**	**0.000**	**-.0588754**	**-.0412514**
_cons	**.0094254**	**.0007558**	**12.47**	**0.000**	**.0079398**	**.010911**
sigma_u	**.00527018**					
sigma_e	**.01475708**					
rho	**.11311445**	(fraction of variance due to u_i)				

Table A3: Breusch-Pagan LM test for cross-sectional dependence (Life Expectancy at birth)

Correlation matrix of residuals:

	__e1	__e2	__e3	__e4	__e5	__e6	__e7	__e8	__e9	__e10	__e11	__e12
__e1	1.0000											
__e2	0.0462	1.0000										
__e3	-0.1428	0.1290	1.0000									
__e4	-0.1821	0.2986	0.4749	1.0000								
__e5	0.0864	0.1968	-0.0570	-0.0430	1.0000							
__e6	0.2410	-0.2508	-0.2351	-0.0370	-0.0159	1.0000						
__e7	-0.1944	0.3217	0.4218	0.3915	0.2232	0.0828	1.0000					
__e8	0.1381	0.3816	-0.0933	0.2949	0.2279	0.1622	0.1349	1.0000				
__e9	-0.0701	0.6362	0.1655	0.1728	0.2123	-0.1525	0.6218	0.1353	1.0000			
__e10	0.0708	-0.4235	0.5377	0.1701	0.0212	0.0302	0.2673	-0.0798	-0.0543	1.0000		
__e11	0.0009	-0.0398	0.2439	-0.0983	0.1531	-0.2787	-0.0113	0.0208	-0.0768	0.3175	1.0000	
__e12	-0.0181	0.4275	0.2325	0.5472	0.0305	-0.1481	0.2045	0.6418	-0.0213	0.1259	0.2994	1.0000
__e13	0.3867	-0.0139	-0.0951	-0.4252	0.1709	0.1894	-0.0955	0.0712	0.0556	0.1597	0.4447	-0.0647
__e14	-0.1387	0.3001	0.3851	0.4439	0.1518	0.0960	0.3996	0.3950	0.2320	-0.0080	0.2470	0.3385
__e15	0.1443	-0.2244	0.4052	0.1211	0.0188	0.0766	0.0758	-0.1137	0.0004	0.2641	-0.1624	-0.3038
__e16	0.0083	0.2290	-0.3998	-0.2000	0.3736	-0.1666	-0.0197	0.4477	0.2855	-0.3334	-0.1703	-0.0338

	__e13	__e14	__e15	__e16
__e13	1.0000			
__e14	0.1261	1.0000		
__e15	-0.1619	0.2656	1.0000	
__e16	-0.0546	-0.0290	0.0150	1.0000

Breusch-Pagan LM test of independence: chi2(120) = **195.782**, Pr = **0.0000**
Based on 26 complete observations over panel units

Table A4: Test for heteroscedasticity (Life Expectancy at birth)

```
Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (16)  =    1664.75
Prob>chi2 =      0.0000
```

Table A5: Breusch-Pagan LM test for cross-sectional dependence (ARD)

Correlation matrix of residuals:

	__e1	__e2	__e3	__e4	__e5	__e6	__e7	__e8	__e9	__e10	__e11	__e12
__e1	1.0000											
__e2	0.0471	1.0000										
__e3	-0.1531	0.1001	1.0000									
__e4	-0.1999	0.2461	0.4154	1.0000								
__e5	0.0984	0.1340	-0.2243	-0.1150	1.0000							
__e6	0.2500	-0.2393	-0.2784	-0.0741	0.0626	1.0000						
__e7	-0.2194	0.2887	0.3662	0.3203	0.2033	-0.0217	1.0000					
__e8	0.1432	0.3741	-0.1289	0.2684	0.2462	0.1911	0.0963	1.0000				
__e9	-0.0748	0.6228	0.1409	0.1208	0.1464	-0.1394	0.6157	0.1394	1.0000			
__e10	0.0854	-0.4873	0.4640	0.1344	-0.0254	0.1024	0.2695	-0.0556	-0.0968	1.0000		
__e11	0.0341	-0.0591	0.2247	-0.1277	0.0945	-0.2762	0.0143	0.0146	-0.0890	0.2467	1.0000	
__e12	-0.0013	0.4406	0.2275	0.5616	-0.0047	-0.1503	0.1857	0.6625	0.0023	0.0982	0.2613	1.0000
__e13	0.4641	-0.0907	-0.2857	-0.5292	0.1490	0.3243	-0.1725	0.1226	0.0346	0.0746	0.3356	-0.1230
__e14	-0.1569	0.2794	0.3248	0.3942	0.1121	0.0711	0.3289	0.3692	0.2330	-0.0317	0.2331	0.3449
__e15	0.1381	-0.2273	0.3789	0.1025	0.0159	0.0770	0.0032	-0.1167	-0.0086	0.2646	-0.1692	-0.2878
__e16	0.0422	0.2521	-0.3665	-0.2434	0.4267	-0.1157	-0.0500	0.4656	0.3234	-0.2791	-0.0905	-0.0042

	__e13	__e14	__e15	__e16
__e13	1.0000			
__e14	0.1064	1.0000		
__e15	-0.1519	0.2064	1.0000	
__e16	0.1398	-0.0044	0.0110	1.0000

Breusch-Pagan LM test of independence: chi2(120) = **190.013**, Pr = **0.0000**
Based on 26 complete observations over panel units

Table A6: Test for heteroscedasticity (ARD)

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (16) = **1890.78**
Prob>chi2 = **0.0000**

Table A7: ARD t-1 and t-2

lGDPLCUdiff	Drisc/Kraay Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lARDdiff	-.0120296	.0387184	-0.31	0.759	-.0919405	.0678814
lagARD	-.0519979	.0321475	-1.62	0.119	-.1183471	.0143513
lag2ARD	.0337085	.0226111	1.49	0.149	-.0129584	.0803755
lFCFdiff	.115919	.0505352	2.29	0.031	.0116194	.2202186
lLabordiff	1.024701	.5417482	1.89	0.071	-.0934121	2.142814
_cons	.0105862	.0185737	0.57	0.574	-.027748	.0489205

Table A8: Life expectancy t-1 and t-2

lGDPLCUdiff	Drisc/Kraay Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lLifeexpDiff	**-7.215012**	**2.81423**	**-2.56**	**0.017**	**-13.0233**	**-1.406726**
lagLifeexp	**11.39909**	**4.835039**	**2.36**	**0.027**	**1.420059**	**21.37812**
lag2Lifeexp	**-4.339844**	**2.331325**	**-1.86**	**0.075**	**-9.151462**	**.471773**
lFCFdiff	**.1025191**	**.0379038**	**2.70**	**0.012**	**.0242894**	**.1807488**
lLabordiff	**-.2012169**	**.297373**	**-0.68**	**0.505**	**-.8149646**	**.4125309**
_cons	**.0468849**	**.0072433**	**6.47**	**0.000**	**.0319354**	**.0618344**

Table A9: Autocorrelation life expectancy

	lLifee~f	lagLif~p	lag2Li~p
lLifeexpDiff	**1.0000**		
lagLifeexp	**0.9330**	**1.0000**	
lag2Lifeexp	**0.7537**	**0.9359**	**1.0000**

Table A10: Autocorrelation ARD

	lARDdiff	lagARD	lag2ARD
lARDdiff	**1.0000**		
lagARD	**0.9474**	**1.0000**	
lag2ARD	**0.9032**	**0.9557**	**1.0000**