On the Effects of Gross Domestic Product on Life Expectancy in Sub-Saharan Africa

Einari Stenberg

School of Science

Bachelor's thesis

Espoo November 7, 2023

Supervisor

Assoc. Prof. Pauliina Ilmonen

Advisor

Assoc. Prof. Pauliina Ilmonen

Aalto University
School of Science

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Tekniikan kandidaatintyön tiivistelmä

Tekijä Einari Stenberg

Työn nimi On the Effects of Gross Domestic Product on Life Expectancy in Sub-Saharan Africa

Koulutusohjelma Teknistieteellinen kandidaattiohjelma

Pääaine Matematiikka ja systeemitieteet **Pääaineen koodi** SCI3029

Vastuuopettaja Assoc. Prof. Pauliina Ilmonen

Työn ohjaaja Assoc. Prof. Pauliina Ilmonen

Päivämäärä November 7, 2023 **Sivumäärä** 21+1 **Kieli** Englanti

Tiivistelmä

Maailmanlaajuinen elinajanodote syntymähetkellä on yli kaksinkertaistunut 120 vuodessa. Elinajanodote on laajasti käytetty mittari, koska sillä voidaan arvioida väestön terveyden kehitystä tietyllä alueella. Elinajanodotteeseen vaikuttavia tekijöitä tutkitaan monista syistä. Yksi näistä syistä voi olla poliittisen päätöksenteon ohjaaminen.

Tässä kandidaatintutkielmassa tarkastellaan Saharan eteläpuolisen Afrikan elinajanodotteen kehitystä. Lisäksi tutkitaan bruttokansantuotteen (BKT) ja elinajanodotteen välistä viiveyhteyttä. Tätä yhteyttä tutkitaan Grangerin kausaalisuustestauksen avulla. Grangerin kausaalisuustesti pyrkii selvittämään, voiko yhden muuttujan kehitys vaikuttaa toisen muuttujan kehitykseen. Testi siis arvioi kahden muuttujan välistä syy-seuraus suhdetta. Tutkielman tavoitteena on selvittää, voiko arviota elinajanodotteen kehityksestä parantaa ottamalla BKT huomioon.

Monet tekijät voivat vaikuttaa elinajanodotteeseen. Vaikuttavia asioita kehittyvissä maissa voivat olla muun muassa lapsikuolleisuus, nälänhätä ja terveydenhuollon tila. Elinajanodotteen kehityksen syitä on tärkeää ymmärtää, sillä elinajanodotteessa esiintyy maailmanlaajuisesti edelleen paljon epätasa-arvoisuutta.

Elinajanodotteen ja BKT:n väliltä ei löytynyt odotettua syy-seuraus suhdetta. Tulosten perusteella BKT ei vaikuta viiveellä elinajanodotteeseen. Viiveyhteys voi silti olla olemassa, mutta elinajanodotteen ja BKT:n välistä suhdetta selittää todennäköisesti jokin monimutkaisempi systeemi. Tulosten perusteella voidaan kuitenkin todeta, että elinajanodotteen kehitys näyttäisi vaikuttavan BKT:n kehitykseen positiivisesti. Laajemmalla tilastollisella tutkimuksella olisi mahdollista parantaa arviota BKT:n ja elinajanodotteen välisestä suhteesta Saharan eteläpuolisessa Afrikassa.

Avainsanat elinajanodote, bruttokansantuote, Grangerin kausaalisuus, Saharan eteläpuolinen Afrikka, aikasarjat

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Einari Stenberg

Title On the Effects of Gross Domestic Product on Life Expectancy in Sub-Saharan Africa

Degree programme Bachelor's Programme in Science and Technology

Major Mathematics and Systems Sciences **Code of major** SCI3029

Teacher in charge Assoc. Prof. Pauliina Ilmonen

Advisor Assoc. Prof. Pauliina Ilmonen

Date November 7, 2023 **Number of pages** 21+1 **Language** English

Abstract

Global life expectancy at birth (LE) has more than doubled in the past 120 years. LE is a widely used measure, because it can be used to assess the health development of a region's population. Factors influencing LE are studied, for example, in order to direct political decision-making.

In this thesis, we investigate the development of LE in Sub-Saharan Africa. Furthermore, we explore the lagged relationship between GDP/capita and LE. This examination is done using Granger causality testing, which aims to determine whether the development of one variable affects the development of another variable. In other words, the test measures causalities between variables. In this study, we aim to see whether the predictability of LE can be enhanced by taking into account GDP/capita.

Several variables influence LE. These factors can include child mortality, famine or the state of healthcare. Understanding the reasons behind changes in LE is crucial as there still exists a lot of inequality in LE globally.

The expected causal relationship was not found between LE and GDP/capita. Based on the results, GDP/capita growth does not cause LE development. It is possible that the causal relationship exists, although the relationship is likely to be more complex than assumed. In contrast to what was assumed, it seems that the development of LE positively influences GDP/capita growth. One explaining reason could be the demographic changes in the population as LE develops. With more comprehensive statistical research, it would be possible to improve the knowledge around the relationship between LE and GDP/capita in Sub-Saharan Africa.

Keywords life expectancy, gross domestic product, Granger causality, Sub-Saharan Africa, prediction, time series

Contents

1 Introduction

United Nations Development Program introduced the Human Development Index in 1990. (The United Nations Development Programme, 1990) The aim was to create a measure that summarizes leading dimensions of human development. These dimensions are a long and healthy life, access to knowledge and a decent standard of living. Among these three, long and healthy life is perhaps the clearest indicator of well being. Expressing long and healthy life is commonly done with an indicator called life expectancy at birth (LE). LE refers to the number of years that individuals born during a specific year would, on average, live if mortality rates remain constant across individual's life. (The World Bank, 2023d) It can be affected by variables such as infant mortality and deaths caused by famine or preventative illness. This is why LE is an excellent measure when estimating the health development of a population.

LE is not only a good indicator of human health. It can also be a predictor of national development. For instance, an increase in LE can indicate that individuals remain productive longer. This obviously increases possibilities to experience life, but the productivity growth can also increase measures of economic performance of a population.

In this thesis we study the effects of Gross domestic product per capita (GDP/capita) on LE in Sub-Saharan Africa (SSA). In particular, we are interested on the lagged effect of GDP/capita on LE. In other words, does there exist a lagged causality, and if there exists, how large is that lag. The hypothesis is that GDP/capita has a strong positive causal relationship on LE. This hypothesis is chosen around the generally accepted positive correlation between GDP/capita and LE. The aim of this thesis is to quantify this causality and determine whether there exists a statistically significant lagged correlation between GDP/capita and LE in SSA.

The chosen area of interest is SSA. We take into account all countries on the African continent that are south of the Saharan desert. In addition, countries that are partly in the Sahara are included. Overall the dataset includes 48 countries, largest by GDP/capita being Seychelles and smallest being Burundi. (The World Bank, 2023a) This area is chosen as it represents a large uniform region of developing countries. We also compare the results to the development of Finland at around the same time span.

The main statistical method in this thesis is Granger causality testing. This method of statistical testing aims to unwrap the underlying time lagged correlations in the data. In this context our goal is to examine, if we can create a time series model with GDP/capita and LE that can predict chances in LE better, than a time series model having LE as its only variable. In practise, this is done with Granger causality testing.

This thesis consists of 7 sections. After the introduction we introduce the data used in this thesis and we reflect on the summary statistics of the dataset. The literature review consist of the current level of research around this topic. After the literature

review we introduce the methodology used in this thesis as well as the results. The discussion consists of critical reflection on the results and the data. Finally, the summary gathers everything established in the thesis into one section.

A tool of general artificial intelligence was used in this study. The AI service used was ChatGPT-3.5 and it was utilised same way as any search engine could have been used. ChatGPT-3.5 was used to find synonyms and all other outputs used from ChatGPT-3.5 were basic instructions on how to create plots and save them with high resolution. Furthermore, everything considered was fact checked from external sources.

2 Data and Variables

GDP/capita is a commonly used measure of a country's economic activity and its growth. It represents the total value of all goods and services produced within a country over a specific period, usually over a year. (The World Bank, 2023b) Overall, GDP serves as an indicator for assessing the economic health, growth, and development of a country. It provides insights into the size of an economy, the level of production and consumption, and changes in economic output over time. This thesis examines GDP growth in current US.

In this thesis, we examine development of LE over time. The data is recovered from the database of World Bank. (The World Bank, 2023c) LE of an individual is affected by the existing conditions in their local environment. In countries that are less developed, LE tends to be comparatively lower than in more developed nations. One explaining reason might be that infant mortality rates are higher in regions that have lower disability adjusted LE. (Reidpath and Allotey, 2003)

The data consists of 62 data points ranging from 1960 to 2021 and the data has yearly intervals. LE represents values as a weighted average from every year and GDP/capita represents gross domestic product divided by midyear population.

For SSA, LE ranged from 41.42 to 61.24 years, with a median of 49.70 years and a mean of 50.43 years. The GDP/capita in SSA varied between 136.72 and 1931.65, with a median of 631.86 and a mean of 796.41. In contrast, Finland had higher LE, ranging from 68.58 to 81.98 years, with a median of 75.02 years and a mean of 75.39 years. The GDP/capita for Finland spanned from 1179.35 to 53772.79, with a median of 22183.43 and a mean of 22859.75. In terms of percentages Finland had GDP growth of 4459.53 % and LE growth of 19.55%. Average growth of GDP/capita has been 7.04 %. Growth of GDP/capita in SSA has in tota been 1312.88 % with an average of 4.68 %. LE in SSA has grown significantly more compared to Finland. The total growth in Finland has been 47.85 %.

Interestingly Finland has had a large GDP/capita precentage growth compared to SSA, but SSA still has had significantly more LE growth. On average, Finland has had 2.36 % points more GDP/capita growth, but LE growth has been 28.3 % points less overall. One direction of speculation could be marginal returns of LE. Meaning

that LE has a biological limit, and as we approach this limit, the development of LE decelerates. Other contributing factors could include the nonlinear decrease of infant mortality over time. (Bishai and Opuni, 2009)

Variable	**Smallest value**	**Largest value**	**Median**	**Mean**
LE, years(SSA)	41.4225	61.2443	49.6983	50.4312
GDP/capita, $(SSA)	136.717	1931.65	631.863	796.414
LE, years(Finland)	68.577	81.9829	75.0203	75.3910
GDP/capita, $(Finland)	1179.3530	53772.7942	22183.4263	22859.74583

Table 1: Summary statistics of SSA and Finland

SSA suffered a big reduction in GDP and a long lasting depression in 1980-2000. There has been a lot of debate on the underlying causes, and it has been generally agreed that there was multiple reasons that caused this depression. For example, it's been shown that the worlds real interest rate as well as the terms of trade have a strong correlation on the economic development in SSA. (Ghura, 1995)

The relationship of LE and GDP/capita seem to be positively correlated. This correlation can be visually analyzed from Figure 1. For example, there is a noticeable decrease in LE after GDP/capita drop in the 1980s and the trend of LE seems to be lower for a significant time period.

In contrast to SSA, Finland has had more stability in it's GDP/capita growth as well as in LE development. This can be seen from comparing Figure 2 and Figure 1. The magnitude of differences in GDP/capita between SSA and Finland can largely be explained with an earlier start of economic growth in Finland. Finland has had quite continuous growth from the 1860s (Hjerppe, 1989).

Here a comparison to Finland is made in order to understand the magnitude of difference in social indicators of SSA compared to the western countries. For example, we can see from Table 1 that largest LE in SSA is over 20 years lower compared to Finland.

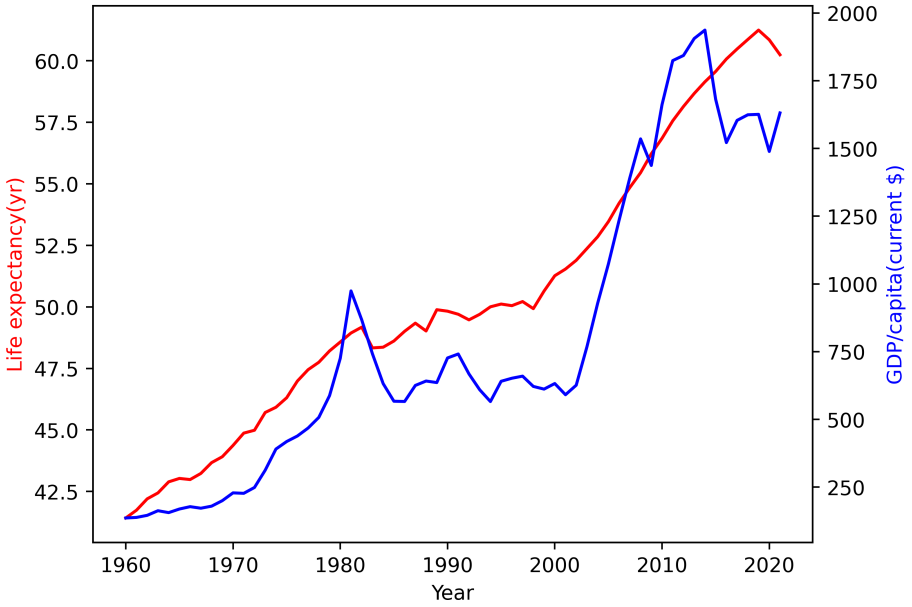

Figure 1: Development of GDP/capita and LE at birth in SSA

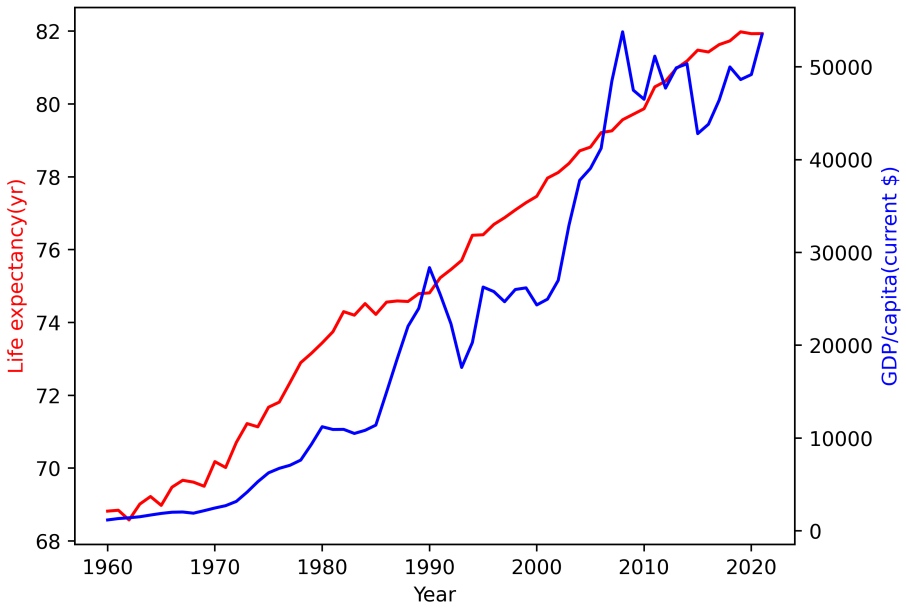

Figure 2: Development of GDP/capita and LE at birth in Finland

It is important to note that here we compare SSA and Finland from the same time period, and not from similar periods of LE growth. This is done as it is important to showcase the magnitude of GDP/capita differences in these regions as well as placing the development of SSA in context of a developed country. Also, from the comparison, we can notice that even in developed countries there are similar trends between GDP/capita and LE.

The data from Finland could've been selected to mirror the same LE growth observed in SSA. Then, similar LE development and the effect of GDP/capita on LE could've be compared. Unfortunately, reliable LE data from that time period is hard to find. One source is Statistics Finland. (Tilastokeskus, 2023a) Statistics Finland has compiled summary data on LE on the decade level. The period during which LE increased from approximately 40 years to 60 years spanned from the 1840s to the 1940s. This could imply that SSA has had more rapid LE growth compared to Finland.

Furthermore, accurate GDP/capita statistics are also quite tricky to estimate from the similar time period. GDP data from Statistics Finland dates only back to 1860s (Tilastokeskus, 2023b) and it has not been adjusted for inflation. Luckily there are some sources that aim to estimate historical GDP/capita growth. According to a study done on Finland's GDP by Riitta Hjerppe (Hjerppe, 1989), the GDP/capita growth at around 1860s to 1940s was on average around 2.1-2.2 %. One can only wonder if the lower growth rate of GDP/capita was the causing reason of lower LE growth at that time. Alternatively, there might have been other explanatory variables at play, such as medicine development or famine. However, this topic is not in scope of this thesis.

3 Literature review

LE is commonly used as an indicator of the well-being of a country or a region. It is usually agreed that the order of causality is that GDP/capita growth causes LE growth. Emanuele Felice et.al (Felice et al., 2016) looked at GDP/capita and LE in Italy and Spain comparing the results to France. They aimed to compare GDP/capita and LE in the long run, spanning 148 years (1861–2008). They saw a non-monotonic relationship between the variables and found that GDP/capita Granger causes LE in certain time periods in Italy and Spain as well as in France. Interestingly, they also found that LE Granger causes GDP/capita. The order of the causality seems to switch as the county or region progresses. The conclusion of this paper was that the causal order of GDP/capita and LE clearly changed during the 148 years in all three countries. It seems to be unclear, how the relationship of causality between LE and GDP/capita behaves in the long run.

There is a possibility that demographic changes in a population affect GDP/capita growth. Chong-Bum An and Seung-Hoon Jeon suggested that there might be an inverted-U shape in the positive correlation measure of LE and GDP/capita. (An and Jeon, 2006) This could imply that rising LE might be beneficial for economic growth when starting from a low LE level, but benefits of this effect start to decrease as the population starts to age.

Lars Kunze aimed to capture the major impacts of rising longevity on economic growth by creating an overlapping-generations model (Kunze, 2014). He concluded that LE impacts economic factors by influencing savings rates, bequests, and public education expenditures. According to Kunze, the balance of these effects will determine whether

LE growth is beneficial to economic growth. In his paper, Kunze did not consider that economic development could have a positive effect on LE.

A substantial volume of literature exist on the causal assumption of GDP/capita growth causing LE growth. Zhiheng Chen et.al (Chen et al., 2021) studied the effects of economic development and environmental factors on LE. They looked at both developed and developing countries. An intercontinental distribution of countries was used from Europe, Asia, Africa and the Americas. One of the major results of the study was the development of a linear regression model aiming to predict LE from variables such as GDP, Gini coefficient and fertilizer consumption. Overall, nine different variables were used in the model. They were able to show that GDP/capita had the greatest impact on LE over the other variables and more importantly, GDP/capita has a strong correlation to LE among developing countries. Interestingly, fertilizer consumption and green house gas emissions had a significant negative effect on LE. Multicollinearity was not considered in the study and this might have caused some errors in the results.

Mayhar Hami studied the relationship between economic growth and LE in Iran during 1966-2013 (Mahyar, 2016). Similarly, Weixiang Luo and Yu Xie conducted a wide study looking at China's trends in income inequality, economic growth and LE (Luo and Xie, 2020). Both of these studies used indicator data from World Bank, which is also the database used in this thesis. Although, Granger causality testing was not used in these studies, they give rigorous results on the causality of GDP/capita and LE.

In Hami's study on Iran's longevity, a vector error correction model was used in order to determine long-run co-integrations between GDP/capita and LE. He found strong correlation with over 94% certainty in the results. Similar results can also be seen in China. Luo and Xie used a regression model with logarithmic GDP, Gini coefficient and a linear term. They found a strong predictable correlation between GDP/capita and LE. The conclusion was that 1 percent growth in GDP/capita would result in an increase of LE of 0.069 years.

4 Methodology

Granger causality is a widely used tool that was first introduced by Clive Granger more than half a century ago. It was developed for analyzing correlations between various time series. (Granger, 1969) The aim is to measure correlations between different phenomena where phenomena Y might be affecting results of phenomena X. Granger causality testing takes into account delayed correlations and this is why Granger causality is seen as a beneficial tool for causality testing.

Granger causality is used in different fields varying from medicine to economics and has even been used in a pollution study (Hoffmann et al., 2005). Furthermore, it has also been shown that Granger causality can prove to be a useful tool in brain imaging (Roebroeck et al., 2005) and even in estimating downside risks between

financial markets (Hong et al.).

Generally Granger causality has been limited to bivariate autoregressive processes due to computational limitations and problem complexity (Shojaie and Fox, 2021). However, recently there has been research advancements applying Granger causality to multivariate time series, called the study of network Granger causality (Basu et al., 2015). In this thesis, we focus on studying a bivariate autoregressive time series model.

Let $X = \{x_t : t \in T\}$ and $Y = \{y_t : t \in T\}$ be stationary time series and let T be an index set of length s. Our data has 62 data points meaning that s=62. Furthermore, let X denote LE and Y denote GDP/capita. In his original paper, Granger proposed that existence of Granger causality can be stated between X and Y, when the variance of a prediction based on X can be reduced by incorporating additional historical information from Y.

Let's determine a bivariate autoregressive model (AR) with n+m lag terms:

$$x_t = \sum_{i=1}^{n} \phi_i x_{t-i} + \sum_{i=1}^{m} \gamma_i y_{t-i} + \epsilon_t, \tag{1}$$

where $\phi = [\phi_1,\phi_2,...,\phi_n]^T$ and $\gamma = [\gamma_1,\gamma_2,...,\gamma_m]^T$ are coefficients of the lag terms and ϵ is the error term. Traditional estimation methods for the coefficients are usually based on maximum likelihood methods and minimizing squares of error terms. (Brockwell and Davis, 1991)

The bivariate Granger causality test contains two steps. First, we create an AR(n) model from X which models and aims to predict LE. This model works as a reduced model in our analysis. The amount of lag terms is determined by analyzing autocorrelations and partial autocorrelations of X.

After we have formed the reduced model AR(n) for X, we form the full Model (1). Y terms are incrementally added to the model and the terms are selected until they no longer improve the predictability of our model. Alternatively, we can predetermine the specific amount of lag terms we want to study.

The predictability is assessed by using a series of t-tests and F-tests. The t-test aims to determine whether individual Y terms have significant coefficients. The F-test is used to test whether Y terms improve prediction accuracy of the reduced model. The t-test is important, when the goal is to build a model, as it is used to evaluate statistical significance of the γ coefficients. However, in the context of this thesis, the t-test does not improve the accuracy of results, as we are not aiming to build a model.

We are using a F-test as the primary statistical test for our null hypothesis. The p-value is calculated by evaluating the F-value. Formally, the F-value can be represented as

$$F = \frac{(RSS_{red} - RSS_{full})/(n+m-n)}{(RSS_{full})/(T-n)},$$

where RSS_{red} and RSS_{full} denote the residual sum of squares for the reduced and full models with m and n parameters. (Shojaie and Fox, 2021) If the calculated test statistic F surpasses the (1 - α)% quantile threshold of an F-distribution with degrees of freedom n+m-n and T - n , it can be stated that there is Granger causality.

Now, the null hyphothesis can be formulated as follows

$$H_0 : [\gamma_1,\gamma_2,...,\gamma_m]^T = [0,0,...,0]^T$$
$$H_1 : [\gamma_1,\gamma_2,...,\gamma_m]^T \neq [0,0,...,0]^T,$$

where, γ_m indicates coefficients of the Model (1). In this thesis we are using $\alpha = 0.05$, as 95% quantile accuracy is enough for the causality confidence we aim to achieve.

4.1 Assumptions

The bivariate time series model works as the framework for identifying Granger causal relationships, but the model relies on a set of implicit and explicit restrictive assumptions that can significantly effect our analysis and results. (Shojaie and Fox, 2021) The most important assumptions in this thesis are the assumption of linearity of causation, and stationarity of the examined time series. Although we limit our investigation to linear causality, it does not rule out the existence of other forms of lagged correlations. Furthermore, before conducting the Granger causality test, it is essential to assess stationarity as our time series are clearly non-stationary.

Other relevant assumptions include the assumption of continuous-valued variables, discrete time and perfectly observed values. We assume that there are no measurement errors in the data. While these assumptions are quite strong, they provide a base for Granger causality testing. All in all, the assumptions should be reviewed when examining data and the results.

5 Results

As there is a clear linear trend in both LE and GDP/capita, we stationarise the time series with first order differencing. After the differencing, both time series seem to be stationarized. This can be seen from Figure 3 and Figure 4.

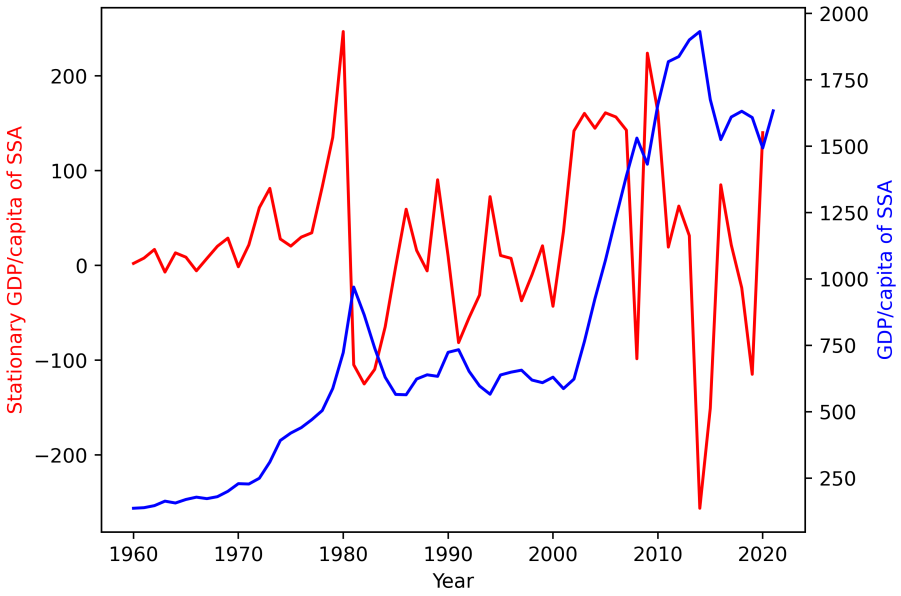

Figure 3: Stationary GDP/capita (red) and GDP/capita development (blue) of SSA

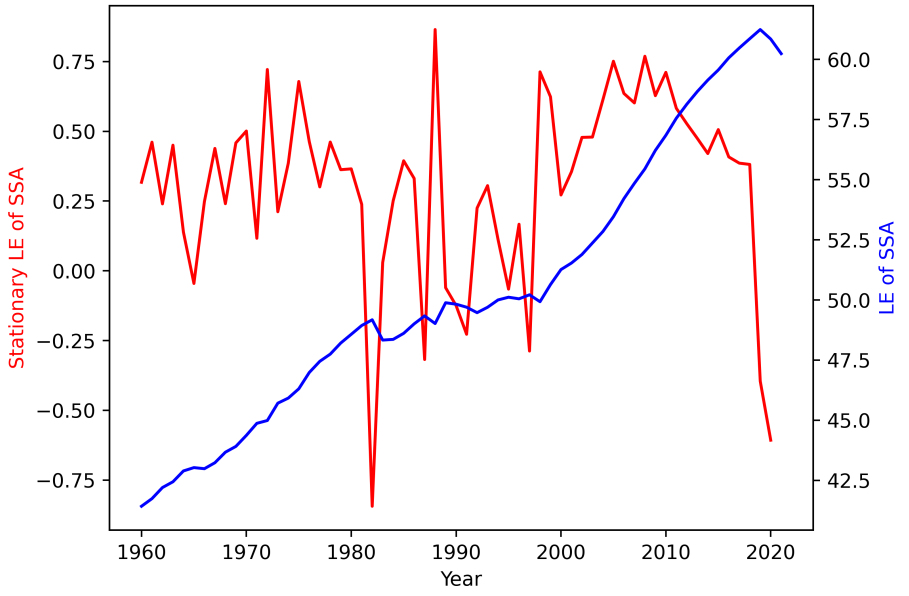

Figure 4: Stationary LE (red) and LE (blue) of SSA

We can determine the existence of correlations inside a time series, by examining autocorrelations (ACF) and partial autocorrelations (PACF). ACF(n) measures overall correlation from time t to t-n and PACF(n) measures the direct correlation

between the observation at time t and at time n. We can see from Figure 5 that after differencing, ACF and PACF decay quickly. After two time steps, there are no significant correlations. This implies that an AR(2) model has enough lag terms for our reduced model.

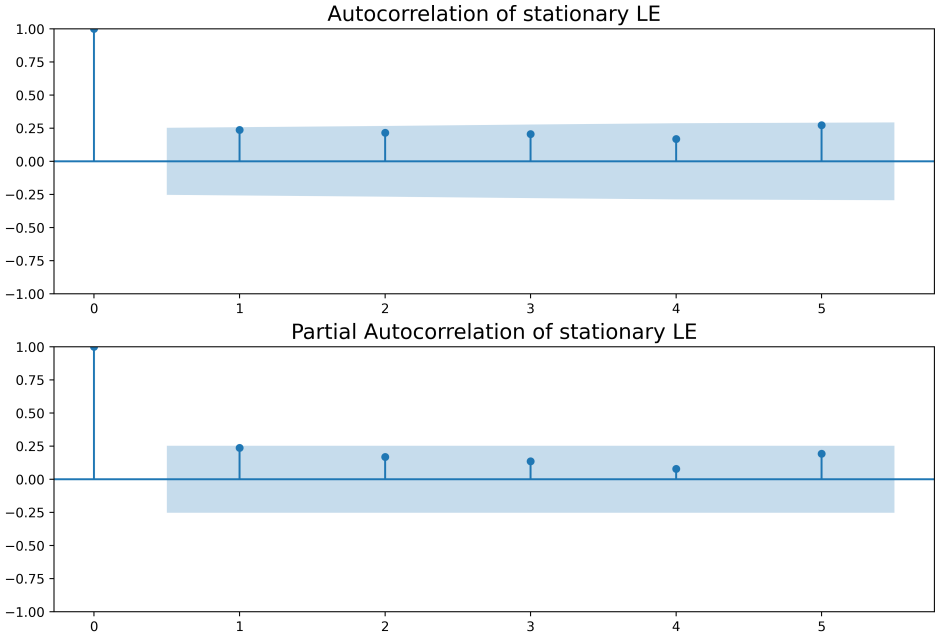

Figure 5: ACF and PACF of stationary LE

We obtain a lot of information about the quality of our stationary time series from graphs, however a more rigorous approach is needed. The augmented Dickey-Fuller (ADF) test is a common way to test for time series stationarity (Dickey and Fuller, 1979). ADF uses a unit root test to determine the existense of a unit root. For both LE and GDP/capita, the p-values of the ADF statistic are very small. The p-value for stationary LE time series ADF test is 0.000002 and the p-value for stationary GDP/capita time series ADF test is 0.000046. Corresponding ADF statistics are spesified in Table 2. Based on the ADF tests, we can reject the null hypothesis of non stationarity and state that neither time series have unit roots. Now we have established a good base for the Granger causality test.

Variable	**ADF statistic**	**p-value**
LE	-5.510939	0.000002
GDP/capita	-4.838330	0.000046

Table 2: Test results of ADF test

It should be noted that there are visible spikes in both Figure 3 and Figure 4. This could be explained with a heavy tailed error term distribution. However, both of the time series look roughly stationary and the ADF tests have small p-values. We can assume that both LE and GDP/capita are stationarized.

It is difficult to interpret lagged causalities heuristically from graphs alone. However, some causality can be interpreted from Figure 6. Generally, when GDP/capita is decreasing, LE has a decreasing trend and vice versa. It is important to note that variances cannot be compared from Figure 6 alone. One reason being that LE and GDP/capita have different scales. However, when analyzing Figure 1, we can see that GDP/capita clearly has more variance around a linear trend.

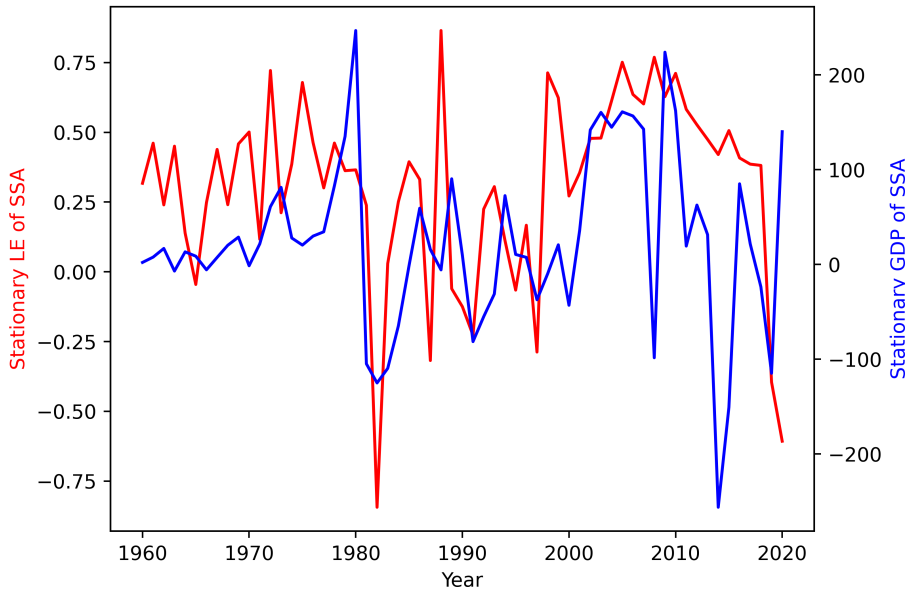

Figure 6: Stationary GDP/capita (blue) and stationary LE (red) of SSA

In this thesis, Granger causality testing is performed with a python package statsmodels. For more about the python package, see (Seabold and Perktold, 2010). This package assumes that n=m, meaning that the full model has the same amount of Y terms and X terms. This assumption doesn't pose any problems as we are not building a model for prediction. We're only testing for the effects of Y on variations of the residuals. From examining the time series as well as PACF we can safely assume that 2 lagged values for the autoregressive models is enough. In addition, the amount of data points used is moderate, thus limiting the reliability of larger time lag evaluations. This means that we'll be examining Granger causality with first and second order lags. Furthermore, as mentioned we choose $\alpha = 0.05$.

Based on the p-values for the Granger causality tests, we cannot reject the null hypothesis. Components of GDP/capita with lags of order one or two don't seem

to increase predictability of an autoregressive LE model. In other words there is no Granger causality between GDP/capita and LE. Table 3 presents the F-values and the corresponding p-values. The p-value for the Granger causality test with first order lag is 0.1664 and the p-value for the Granger causality test with second order lag is 0.4641. Neither surpasses the 95% quantile threshold of an F-distribution.

Num. Lags	F-value	p-value
1	1.9656	0.1664
2	0.7789	0.4641

Table 3: F-values and p-values for the Granger causality test between GDP/capita and LE with first and second order lags

Interestingly, we found that LE Granger causes GDP/capita. The p-value for lag of one year was 0.0220 and for lag of two years 0.0484. Corresponding F-values were 5.5533 and 3.2088. The results for one year of lag is the most significant. These results would imply that increase in LE has a positive effect on economic development in SSA.

6 Discussion

In this thesis we looked into time delays of one and two years. With these lags, GDP/capita does not Granger cause LE in SSA. It is difficult to say whether significant causalities with longer time delays could have been found. The effect that GDP/capita has on LE could have a larger delay, however based on the Figure 1, it seems that this should not be the case.

It is possible that the causality could have been found for some other time period. As seen in Italy and Spain (Felice et al., 2016) the order of causality can flip when a region gains a certain level of economic development. This also could have happened in SSA and it would explain the direction of causality we observed.

Causal linearity was assumed and it was not tested. This opens a possibility for the existence of a nonlinear causal relationship between LE and GDP/capita. More over, it is likely that there are other variables that have strong correlation with LE which could offer a better explanation for the development of LE. These could include investments in public healthcare or political decisions.

The results suggest a one directional causal relationship between LE and GDP/capita. It seems that increase in LE has a positive effect on GDP/capita. One explaining reason could be longer life spans. Literature seems to suggest that longer life spans have a significant effect on the GDP/capita growth. (An and Jeon, 2006) This could be an origin of the observed causality, although generally, in developing countries, majority of early LE growth comes from decrease in infant mortality. Thus it's difficult to say whether longevity is the driving factor of the observed causality. Furthermore, it is not within the scope of this thesis to evaluate whether the average

years spent in adulthood has actually increased in SSA. It is possible that the increase in LE we see in this thesis, comes almost entirely from the decrease of infant mortality rates. This is left for future researchers to study.

It appears that there is no single rule dictating the causality relationship between GDP/capita and LE. Instead, we have demonstrated that quantifying causalities is complicated and usually involves more than one variable to capture a certain phenomena. For example, it has been shown that fertility rates and daily calorie intakes have a strong correlation with LE in developing countries. (Husain, 2002) These variables don't necessarily have a causal relationship on GDP/capita. More research is needed, if a strong understanding on causalities affecting LE in SSA is to be established.

There could also have been some errors in the data or methods. The data was collected from a commonly used database: The World Bank. Thus, the data we used in this thesis was most likely accurate and reliable. Nevertheless, it is possible that there was some measuring errors for example in the LE dataset.

7 Summary

The research question of this thesis was: does there exist lagged causality between Life expectancy (LE) and Gross domestic product per capita (GDP/capita) in Sub-Saharan Africa (SSA). The main objective was to quantify this possible causal relationship and to determine the amount of statistically significant lag terms. This was done by introducing Granger causality testing. Our hypothesis was that there exists a statistically significant causal relationship between LE and GDP/capita with a certain time lag in SSA. The hypothesised direction of causality was that changes in GDP/capita would have an effect on LE. A lot of research has been done around the relationship of GDP/capita and LE. However, there is a limited body of literature specifically investigating the Granger causality relationship between LE and GDP/capita in SSA.

It seems that our hypothesised direction of causality was incorrect. No significant causality could be found in the hypothesised causal order. However, we found that causality seemed to be in the other direction. We found that LE is Granger causing GDP/capita. More precisely, LE development has a positive effect on GDP/capita with lags of one and two years. Additional lags were not explored due to uncertainty of accuracy.

Existing literature is in line with our result. Research generally agrees that there is some causal relationship between LE and GDP/capita, although there is no clear consensus on the direction of the causality. It is possible that there exists a positive relationship between GDP/capita and LE in SSA, but based on the results of this thesis, this possibility cannot be confirmed.

Our findings were partly unexpected. We did not find causality in the suspected direction, however the results suggest that LE is Granger causing GDP/capita. This

would mean that as LE continues to rise, there should be a corresponding increase in GDP/capita in SSA. Following research can also be extended to predicting the established relationship using, and extending the methods applied in this thesis. However, wider investigation is needed around the other potential causal variables. This should be done before using the established results as a basis of decisions. This thesis functions as a good framework for future studies on LE and GDP/capita in SSA.

We have provided insights into the relationship between GDP/capita and LE. The results of this thesis can be extended into general knowledge about the relationships of GDP/capita and LE, but not without considering the assumptions. In reality it is likely that there exists a more complex system underneath explaining the relationship between LE and GDP/capita. It is left for further research to consider this possibility.

References

Chong-Bum An and Seung-Hoon Jeon. Demographic change and economic growth: An inverted-u shape relationship. *Economics Letters*, 92(3):447–454, Sep 2006. URL https://www.sciencedirect.com/science/article/pii/S0165176506001170. (accessed 1.8.2023).

Sumanta Basu, Ali Shojaie, and George Michailidis. Network granger causality with inherent grouping structure. *The Journal of Machine Learning Research*, 16(1):417–453, 2015. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8278320. (accessed 7.10.2023).

David Bishai and Marjorie Opuni. Are infant mortality rate declines exponential? the general pattern of 20th century infant mortality rate decline. *Population Health Metrics*, 7(1):13, 2009. URL https://doi.org/10.1186/1478-7954-7-13. (accessed 23.8.2023).

Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*, pages 238–272. Springer Science Business Media, 2th edition, 1991.

Zhiheng Chen, Yuting Ma, Junyi Hua, Yuanhong Wang, and Hongpeng Guo. Impacts from economic development and environmental factors on life expectancy: A comparative study based on data from both developed and developing countries from 2004 to 2016. *International Journal of Environmental Research and Public Health*, 18(16):8559, 2021. URL https://www.mdpi.com/1660-4601/18/16/8559. (accessed 15.8.2023).

David A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74 (366):427–431, 1979. URL https://www.jstor.org/stable/2286348. (accessed 19.9.2023).

Emanuele Felice, Josep P. Andreu, and Carlo D'Ippoliti. GDP and life expectancy in italy and spain over the long run: A time-series approach. *Demographic Research*, 35:813–866, 2016. URL https://www.jstor.org/stable/26332096. (accessed 4.10.2023).

Dhaneshwar Ghura. Macro policies, external forces, and economic growth in sub-saharan africa. *Economic Development and Cultural Change*, 43 (4):759–778, 1995. URL https://www.jstor.org/stable/pdf/1154556.pdf?casa_token=g0xfZTKvROgAAAAA:QRf0HLP9n_C5TYlo6MWVPiL0BT4_cvkwdDApJ_xj-atlFJfgXa3QdFnWTMnvtO_iJaLauMYB6dSiBM_1xO80ShzGG_Rn9jae8-heFWt1egkiODWMVIHx. (accessed 12.8.2023).

Clive W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969. URL http://www.jstor.org/stable/1912791. (accessed 1.7.2023).

Riitta Hjerppe. The finnish economy 1860-1985 : Growth and structural change. *Suomen Pankki*, pages 41–60, 1989. URL https://publications.bof.fi/handle/10024/44762. (accessed 1.8.2023).

Robert Hoffmann, Chew-Ging Lee, Bala Ramasamy, and Matthew Yeung. FDI and pollution: a granger causality test using panel data. *Journal of International Development*, 17(3):311–317, 2005. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jid.1196. (accessed 10.8.2023).

Yongmiao Hong, Yanhui Liu, and Shouyang Wang. Granger causality in risk and detection of extreme risk spillover between financial markets. *Journal of Econometrics*, 150(2):271–287. URL https://www.sciencedirect.com/science/article/pii/S0304407608002248. (accessed 4.10.2023).

Abhar R. Husain. Life expectancy in developing countries: A cross-section analysis. *The Bangladesh Development Studies*, 28(1/2):161–178, 2002. URL http://www.jstor.org/stable/40795653. (accessed 22.10.2023).

Lars Kunze. Life expectancy and economic growth. *Journal of Macroeconomics*, 39: 54–65, 2014. URL https://www.sciencedirect.com/science/article/pii/S0164070414000032. (accessed 29.8.2023).

Weixiang Luo and Yu Xie. Economic growth, income inequality and life expectancy in china. *Social Science & Medicine*, 256:4–8, 2020. URL https://www.sciencedirect.com/science/article/pii/S0277953620302653. (accessed 29.8.2023).

Hami Mahyar. Economic growth and life expectancy: The case of iran. *Studies in Business and Economics*, 11(1):80–87, 2016. URL https://sciendo.com/article/10.1515/sbe-2016-0007. (accessed 29.8.2023).

Daniel D. Reidpath and Pascale Allotey. Infant mortality rate as an indicator of population health. *Journal of Epidemiology & Community Health*, 57(5):344–346, 2003. URL https://jech.bmj.com/content/57/5/344. (accessed 7.10.2023).

Alard Roebroeck, Elia Formisano, and Rainer Goebel. Mapping directed influence over the brain using granger causality and fMRI. *NeuroImage*, 25(1): 230–242, 2005. URL https://www.sciencedirect.com/science/article/pii/S1053811904006688. (accessed 10.8.2023).

Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010. URL https://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf. (accessed 10.8.2023).

Ali Shojaie and Emily B. Fox. Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, 9(1):289–319, 2021. URL https://doi.org/10.1146/annurev-statistics-040120-010930. (accessed 26.8.2023).

The United Nations Development Programme. Human development report 1990. *Oxford Univ. Pr*, 1990. URL https://hdr.undp.org/system/files/documents/hdr1990encompletenostatspdf.pdf. (accessed 5.8.2023).

The World Bank. Gdp per capita(current us dollars)[data file], 2023a. Data retrieved from World Development Indicators https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.PCAP.CD&country=SSF (accessed 1.8.2023).

The World Bank. Gdp per capita(current us dollars), 2023b. The world Bank: Metadata Glossary, https://databank.worldbank.org/reports.aspx?source=2&type=metadata&series=NY.GDP.PCAP.CD (accessed 1.8.2023).

The World Bank. Life expectancy at birth(years)[data file], 2023c. Data retrieved from World Development Indicators, https://databank.worldbank.org/reports.aspx?source=2&series=SP.DYN.LE00.IN&country= (accessed 1.8.2023).

The World Bank. Life expectancy at birth(years), 2023d. The world Bank: Metadata Glossary, https://databank.worldbank.org/metadataglossary/world-development-indicators/series/SP.DYN.LE00.IN#:~:text=Life%20expectancy%20at%20birth%20indicates,1)%20United%20Nations%20Population%20Division. (accessed 1.8.2023).

Tilastokeskus. Infant life expectancy by gender, 1751-2022, 2023a. Data retrieved from Tilastokeskus Database, https://pxdata.stat.fi:443/PxWeb/sq/c9af09cb-ff9f-421a-9482-56f598188031 (accessed 1.8.2023).

Tilastokeskus. Historical series of national accounts, annually, 1860-2022, 2023b. Data retrieved from Tilastokeskus Database, https://pxdata.stat.fi:443/PxWeb/sq/ee329c46-9375-4648-8407-a4961b70e8da (accessed 1.8.2023).