Linear Regression Analysis On The Impact of Defence Expenditure on Human Capital Development in U.S.A. : A Time Series Econometrics Study.

INTRODUCTION

The main task of any effective government is the safety of people and their properties, which can be done through continuous investments in defense and military forces. The budget for the army of any country shows how much it cares about security and how important it is for its national policy. For many years, the USA has been the leader in military spending in the world and it constitutes about 37% of total military spending and in 2024 it reached approximately $997 billion, which accounts for about 3.4% of GDP. In times of peace after the Cold War era, the main events that influenced US defense spending were the Gulf War, the war in Afghanistan and the war in Iraq. These wars caused the increase in military budget, which increased sharply from approximately $413 billion in 2001 to $833 billion in 2011 because of the operations after September 11.

From 1974 to present day, the US has conducted many military interventions which include Lebanon intervention, Invasion of Grenada, Invasion of Panama, Somalia Intervention, Bosnian War Intervention, Kosovo War, and Libya intervention. Research reveals that military involvement by the US has been quite significant with more than 250 military interventions since 1991. Global military spending hit an all-time high of $2.7 trillion in 2024 amid rising geopolitical tensions and increasing defense spending in various countries. These developments illustrate the strong relationship between geopolitical disputes and increased military spending.

Defense expenditures have both positive and negative effects on the economy and economic development. On one hand, they increase security and can lead to technological innovation; however, on the other hand, they might deprive essential sectors of funds like health and welfare. These are some of the important aspects of human capital development that are taken into consideration in this paper. They are considered under the proxy measure of life expectancy. Although the US economy is very strong, and public expenditure is high, the life expectancy of the country stands at around 78 years, which is quite low in comparison to other developed countries.

OVERVIEW OF THE STUDY

In an attempt to find out whether continuous military spending helps or hinders long term socio-economic development, this paper looks into the effects of military expenditure on human capital development. In order to understand better the issue at hand, this research utilizes statistical analysis, in particular, Ordinary Least Squares Method. The variables of concern here are: Life Expectancy at Birth (Total for the United States; no. of years) (Proxy for Human Capital Development); (1) National Defence Consumption Expenditures and Gross Investment (Billions of Dollars); (2) Government Total Revenue (Percent of Gross Domestic Product); (3) Population Growth Rate (Percent Change at Annual Rate); (4) Gross Fixed Capital Formation (trillions of Dollars).

This paper makes use of the Time Series method. It enables to analyze data collected during different periods of time, thus allowing for identification of the relationship between dependent and independent variables. The variables are extracted as secondary data from 1974-2024, Annually (51 observations).Taking into account the complex nature of this project, it is aimed to use statistical methods and data analysis in order to:

1) Find out what determines human capital development measured by Life Expectancy at Birth in the United States.

2) Identify and quantify the relationship between life expectancy and macroeconomic variables with particular attention paid to military expenditures.

Military expenditures are a part of governmental policy and are always justified in terms of national security. At the same time, any military expenditures have their cost in terms of opportunity cost – in case when the government spends a lot of money on national defence, there is not enough money left to spend on such things as healthcare, education, infrastructure. Although high levels of defence expenditures indicate economic prosperity in some cases, it may happen that it means just less funding for sectors which can influence human development indirectly and in complicated manner.

The results of this project will be highly useful for the policymakers, economists and development experts, who will be able to see the whole picture of defense expenditure and its implications for human development.

MOTIVATION

Base Papers for reference:

1) *"Military Expenditure and Human Capital Development in Nigeria" by Obasi C.N.* , Asogwa F. O., Nwafee F. I. from Department of Economics, University of Nigeria, Nsukka, Nigeria (2018)*

Military spending has been widely recognized as a factor that leads to a crowding out effect in terms of investing in human capital. This paper seeks to determine the effect of military spending on the development of human capital in Nigeria, using yearly data from 1970 to 2014. The findings reveal that military spending has no significant effect on education (gross school enrollment) but has a significant effect on health (infant mortality). Hence, military spending has negative implications for human capital development.

2) *"Military Expenditures And Health Outcomes: A Global Perspective" by Seemab Gillani, Muhammad Nouman Shafiq, Tusawar Iftikhar Ahmad (2019)*

In this research, cross-country panel data from 156 nations (1970-2014) are used to investigate the connection between the military expenditure of a country and its health status in terms of the country's life expectancy. The effect of the military expenditure on health capacity and life expectancy is not always negative. Rather, the connection between military expenditure and health capacity is conditional. Military expenditure is positively correlated with increased health capacity in certain circumstances.

3)*"Military Expenditure and Economic Growth: A Meta-Analysis" by Aynur Alptekin from School of Economics, University of Surrey, Surrey, UK (2012)*

The meta-analysis includes 32 studies analyzing the link between military expenditure and economic growth. The results of the analysis reveal that the effect of defense expenditure on economic growth differs from one country to another. Specifically, military expenditure may lead to economic growth in developed nations but can be insignificant elsewhere. Given that economic growth is strongly associated with increases in the level of income, health care, and public service provision, the meta-analysis indicates how military expenditure might indirectly affect human development outcomes such as life expectancy.

These papers gave me the motivation to find out more about this field and it made me think whether I would want to challenge or study it for a developed country such as the United States. The researchers used different independent variables that helped me identify the right ones for this study.

VARIABLES

1) **Dependent Variable:**

- **Life Expectancy at Birth, Total (Years) in the United States of America (Proxy for Human Capital Development)**: This metric obtained from the Federal Reserve Economic Data (FRED) database, represents the average number of years a newborn is expected to live assuming current mortality rates remain constant throughout their lifetime. It reflects overall population health and captures the combined effects of healthcare access, living standards and socio-economic conditions. **Life expectancy** is widely **used as a proxy** for **human capital development** because it represents a key component of human capital **health alongside education and skills**. Higher life expectancy indicates better healthcare, nutrition, and living conditions, which enhance individuals' productivity, longevity in the workforce, and capacity to invest in skills. In this analysis, it serves as the dependent variable, representing changes in human development and overall well-being. *(Source: FRED)*

2) **Independent Variables:**

- **Federal Government: National Defence Consumption Expenditures and Gross Investment (Billions of Dollars):** This metric, obtained from the Federal Reserve Economic Data (FRED) database measures total federal government spending on national defence including both consumption expenditures and gross investment expressed in billions of dollars. It reflects the scale of military-related economic activity and resource allocation toward defence. In this analysis, it serves as a key independent variable to examine the impact of defence spending on life expectancy *(Source: FRED)*

- **Federal Receipts as Percent of Gross Domestic Product (Percent of GDP):** This variable, sourced from FRED, represents total federal government receipts (including taxes and other revenues) expressed as a percentage of GDP. It indicates the government's revenue-generating capacity relative to the size of the economy. In this study, it is included as an independent variable to capture the role of fiscal capacity in influencing life expectancy. (Source: FRED)

- **Population Growth for the United States (Percent Change at Annual Rate):** This metric, obtained from FRED, measures the annual percentage change in the total population of the United States. It reflects demographic dynamics that can influence resource allocation, healthcare demand, and overall economic conditions. In this analysis, it is used as an independent variable to assess how population changes affect life expectancy. *(Source: FRED)*

- **Gross Fixed Capital Formation in the United States (Trillions of Dollars):** This variable, sourced from FRED, represents the total value of investment in fixed assets such as infrastructure, machinery, and buildings. It serves as an indicator of long-term economic development and productive capacity. In this study, it is included as an independent variable to evaluate the influence of investment on life expectancy. (Source: FRED)

METHODOLOGY

What is Regression Analysis?

Regression analysis is a statistical technique used to explore the relationship between a dependent variable and one or more independent variables. It is commonly used to model the relationship between variables and to make predictions based on that relationship. In a simple linear regression analysis, the relationship between two variables is modelled as a straight line, where the dependent variable is predicted based on the independent variable. In multiple regression analysis, several independent variables are used to predict the dependent variable.

The most widely used method of linear regression is known as the method of ordinary least squares, wherein the regression analysis produces a unique line that minimises the sum of squared differences/residuals between actual/observed values of the dependent variable and the estimated values of the dependent variable (obtained as output from the hypothesised model). The closer the actual values are to the regression line, the better fit the model proves to be.

Why Regression Analysis?

Here are several reasons why regression analysis is a particularly suitable choice for the project investigating factors affecting the human capital development index proxied by life expectancy:

1. **Identifying Relationships and Quantifying Effects**: Regression analysis excels at uncovering the relationships between variables and quantifying their influence. In our case, it allows us to assess how changes in national defence consumption expenditures and gross investment, government total revenue, population growth rate, gross fixed capital formation are statistically associated with variations in human capital development proxied by life expectancy at birth. By estimating the coefficients (β), we could determine the average magnitude and direction of these effects (positive or negative).

2. **Handling Multiple Independent Variables**: Unlike simpler techniques that only analyse one independent variable at a time, regression analysis can handle multiple independent variables simultaneously. This is crucial for our study, as we're interested in the combined influence of several factors on the human capital development proxied by life expectancy at birth. Regression helps us to isolate the effect of each variable while controlling for the effects of the others.

3. **Making Predictions**: Once we've estimated the regression equation, we can use it to make predictions about the life expectancy at birth based on specific values

of the independent variables. This could help estimate how changes in defence expenditures might affect future human capital development proxied by life expectancy **.**

4. **Testing Hypotheses:** Regression analysis facilitates the testing of specific hypotheses about the relationships between variables. For instance, we might hypothesise that a rise in defence expenditures leads to an increase in the human capital development proxied by life expectancy (positive β_1). The regression analysis will provide statistical evidence to support or refute such hypotheses.

5. **The Goodness of Fit:** Regression analysis provides a measure of how well the model fits the data. This can help assess the model's reliability and accuracy and determine whether further analysis is needed.

Overall, regression analysis is a powerful tool for untangling the complex interplay of factors that shape the human capital development proxied by life expectancy in the United States. It allows us to move beyond simple correlations and establish statistically robust relationships to gain a deeper understanding of this crucial aspect of the human development.

The relationship between two variables is modelled by a **simple linear regression** equation which is, $Y_i = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} X_i + \boldsymbol{\epsilon}$ where, Y is the dependent variable, the variable which depends on another independent variable.

- $\boldsymbol{\beta_0}$***(beta-nought)***: This is the intercept, representing the average value of Y when all the independent variables (X) are equal to zero (often not a realistic scenario, but it provides a reference point).

- $\boldsymbol{\beta_1}$***(beta-one)***: This is the slope coefficient, indicating the average change in Y for a one-unit increase in X, holding all other independent variables constant.

- ***ε(epsilon)***: This represents the error term, capturing the random, unexplained variation in Y that is not accounted for by the independent variables.

The goal of the simple linear regression analysis is to estimate the values of β_0 and β_1 that best fit the data and also minimise the error term. Once the values of β_0 and β_1 are estimated, we can use the equation to predict the value of Y for any given value of X.

In a **multiple regression analysis**, there are more than one independent variable and they are collectively used to predict the value of the dependent variable.
The equation is as follows- $Y_i = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} X_{2i} + \boldsymbol{\beta_2} X_{3i} + \dots + \boldsymbol{\beta}_n X_{ni} + \boldsymbol{\epsilon}$, where X_{2i}, X_{3i}, ...,X_{ni} are the n independent variables. The goal is to estimate the 2_i, 3_i, n_i values

of , , , ..., $\beta_0, \beta_1, \beta_2, \beta_n$ that best fit the data and minimise the error term, so that we can use the equation to make predictions based on the values of the independent variables. **GretL,** the statistical software used to run the tests that are important for a regression analysis.

Note: All the tests in the model are conducted at a 5% level of significance.

SUMMARY STATISTICS

The data in the image shows the **summary statistics** of the dependent variable and all the independent variables used in the data. It is used to study the basic statistical terms of a variable.

	Mean	Median	Minimum	Maximum
NationalDefenceE~	513.37	406.46	109.89	1082.7
FederalReceiptsa~	17.050	17.058	14.371	19.756
PopulationGrowth	0.92768	0.92780	0.15732	1.3869
GrossFixedCapita~	2.6430	2.7922	0.92710	5.1000
LifeExpectancyat~	76.282	76.580	71.956	79.000

	Std. Dev.	C.V.	Skewness	Ex. kurtosis
NationalDefenceE~	265.25	0.51669	0.28282	−1.0511
FederalReceiptsa~	1.1584	0.067938	−0.13689	0.072229
PopulationGrowth	0.22588	0.24349	−0.75569	1.8239
GrossFixedCapita~	1.1996	0.45389	0.29072	−1.1025
LifeExpectancyat~	1.9690	0.025812	−0.30442	−1.0170

	5% perc.	95% perc.	IQ range	Missing obs.
NationalDefenceE~	122.16	959.35	397.98	0
FederalReceiptsa~	14.675	18.984	1.1884	0
PopulationGrowth	0.47244	1.3257	0.22808	0
GrossFixedCapita~	1.0010	4.7181	1.8475	0
LifeExpectancyat~	72.756	78.809	3.7707	0

Summary Statistics, using the observations 1974 to 2024

The summary statistics helped us to find the Mean, Median, Minimum, Maximum, Standard Deviation, Coefficient of Variation, Skewness, Ex. Kurtosis, 5% percentile, 95% Percentile, Interquartile Range and Missing Observations of all the variables used in the regression analysis.

COEFFICIENT OF VARIATION

The **Coefficient of Variation (CV**) is a measure of relative variability or dispersion of a variable, expressed as a percentage of the meanX. It is calculated by dividing the standard deviation of the variable by the mean of the variable and then multiplying by 100 to express it as a percentage.

Variability refers to the extent to which data points in a dataset or a variable deviate or differ from each other. It is a measure of how spread out or dispersed the values of a variable are and it quantifies the degree of diversity or heterogeneity in the data. It is typically assessed using measures such as the range, variance, standard deviation, and coefficient of variation (CV). A larger variability indicates that the values of the variable are more spread out or dispersed, having greater uncertainty while a smaller variability indicates that the values are more closely packed or clustered together, meaning more stable and predictable.

	C.V.
NationalDefenceE~	0.51669
FederalReceiptsa~	0.067938
PopulationGrowth	0.24349
GrossFixedCapita~	0.45389
LifeExpectancyat~	0.025812

Interpretation of Coefficient of Variation:

- **Lower CV (less than 10%):** The data points are tightly clustered around the mean, indicating **low variability**.
- **Moderate CV (between 10% and 30%):** The data exhibits **some variability** but isn't excessively high.
- **Higher CV (greater than 30%):** The data is significantly spread out, suggesting **high variability** relative to the mean.

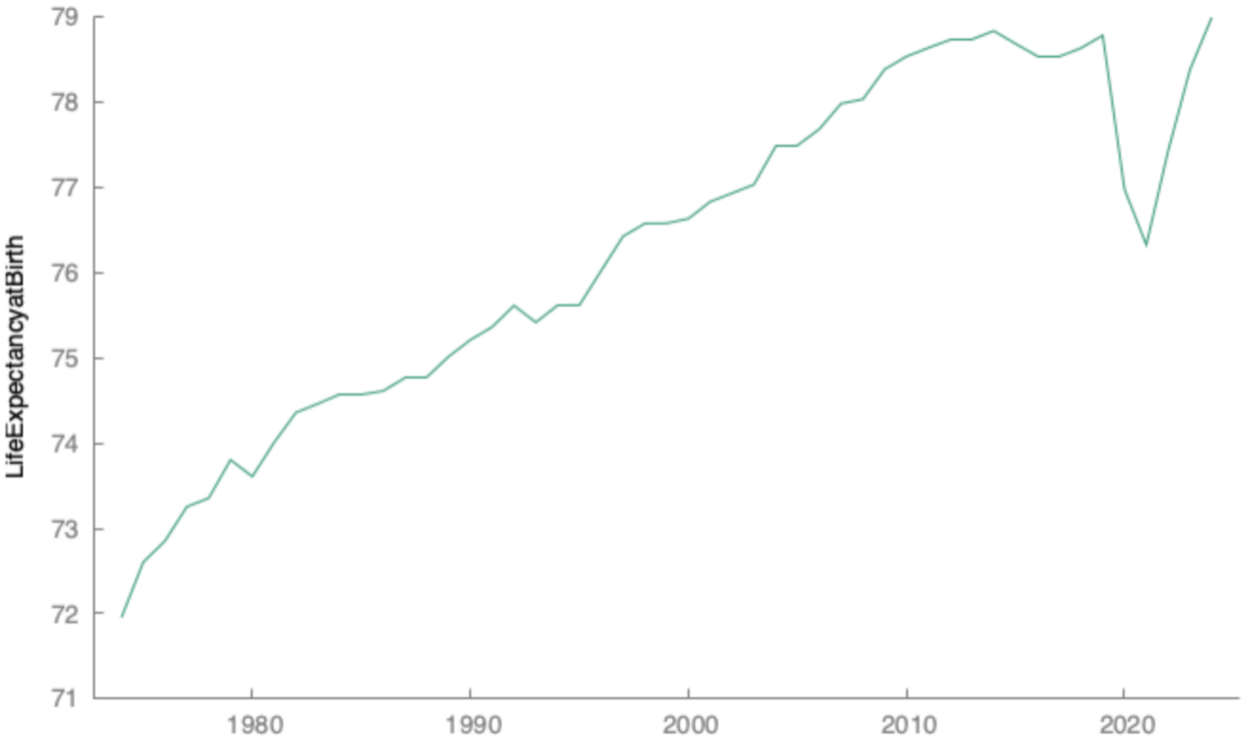
In this data,

1. The coefficient of variation of the **Federal Receipts as Percentage of GDP** and **Life Expectancy at Birth** has low variability since its CV value is *less than 10%.*
2. The coefficient of variation of the **Population Growth (Percent Change at Annual Rate)** have moderate variability since they lie *between 10% to 30%* CV value.
3. The coefficients of variation of the **National Defence Expenditures (Billions of Dollars)** and **Gross Fixed Capital Formation (Trillions of Dollars)** has high variability since its CV value is just *above 30%*.

TIME SERIES ANALYSIS

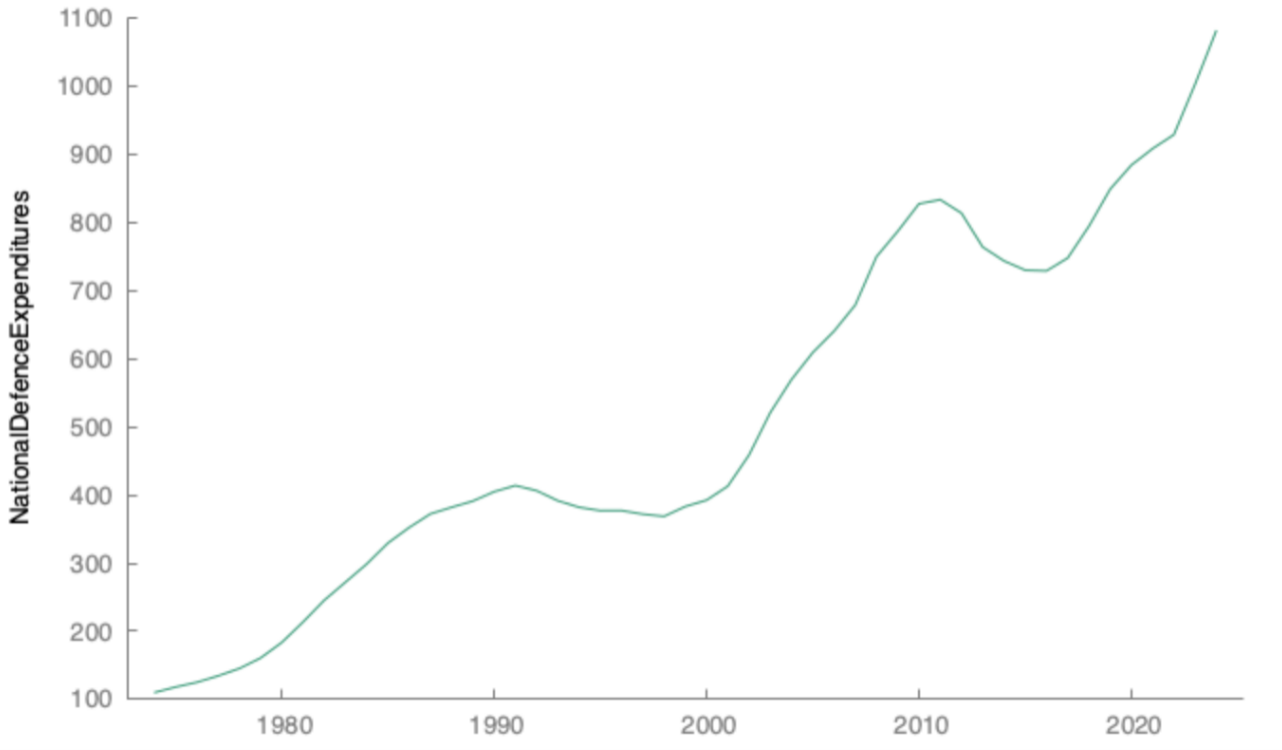
Time series analysis is a statistical tool used to analyse any given data that is collected over a period of time. It delves into understanding the

1. Nature of the sequence of data.
2. Identifying the data's repeated behaviour.
3. Long term trends in the data.



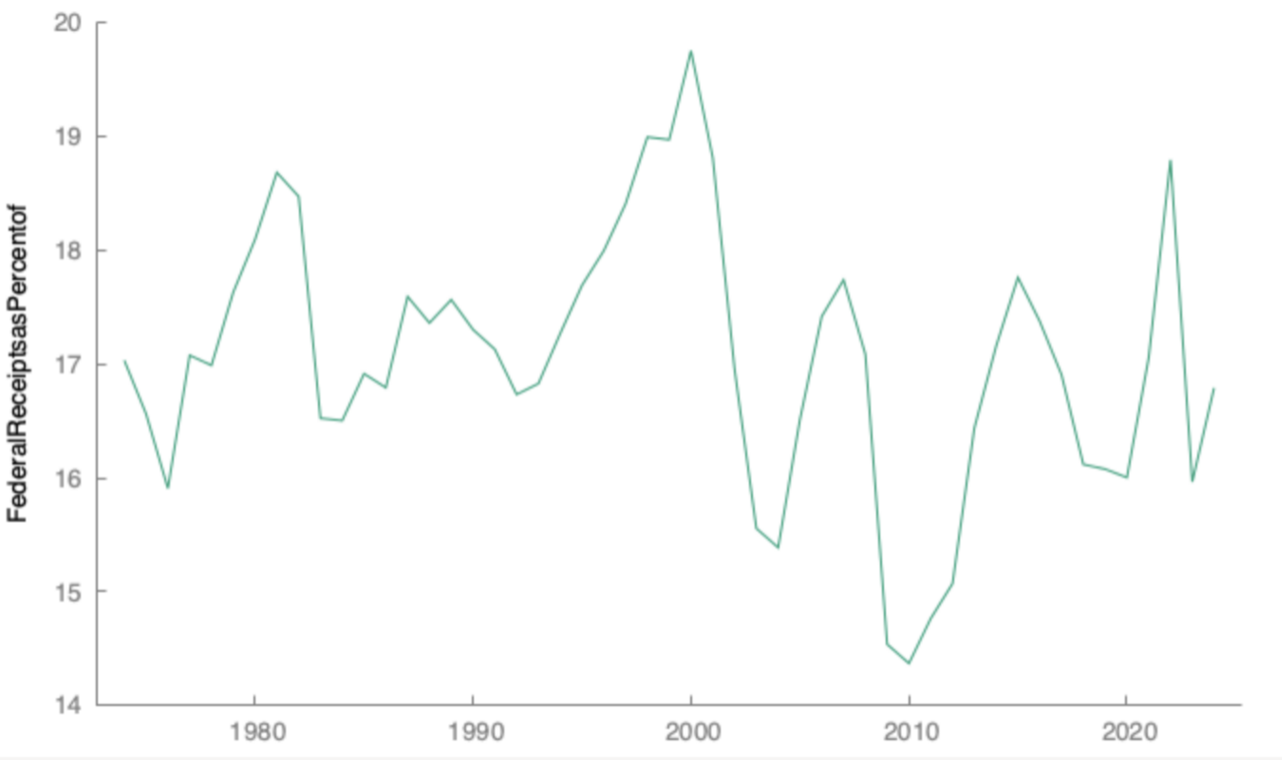
This is the time series plot of **Life Expectancy at Birth** with time in years. **Life Expectancy at Birth** is taken on Y-axis whereas time is taken on X-axis.

The graph shows a steady increase in life expectancy from the late 1970s up to around 2019, indicating continuous improvements over time. There are minor fluctuations in the early 1990s but the overall trend remains upward. Around 2020, there is a sharp decline after which the value rebounds quickly showing a strong recovery. Post-2021, the trend resumes its upward movement reaching its highest level by the end of the period.



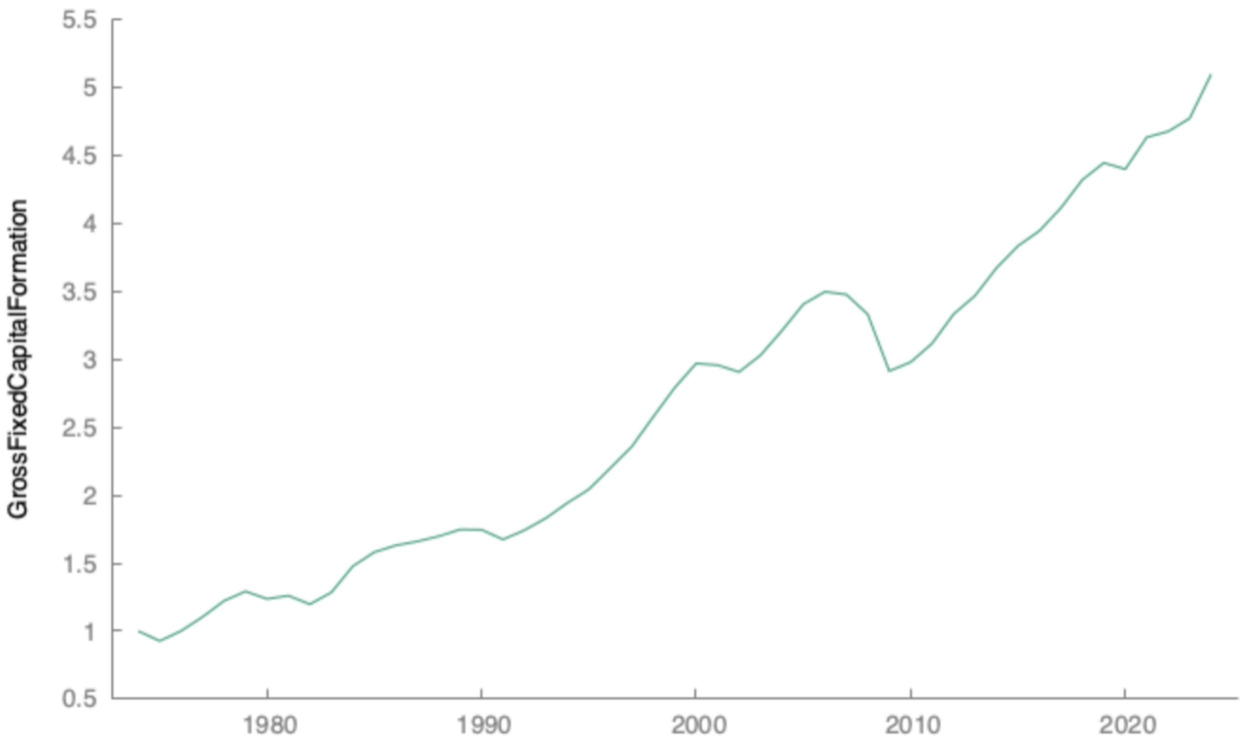
This is the time series plot of **National Defence Expenditures (Billions of Dollars)** with time in years. **National Defence Expenditures (Billions of Dollars)** is taken on Y-axis whereas time is taken on X-axis.

The graph shows defence spending rising steadily in the early years with strong growth through the 1980s and a peak in the early 1990s. This is followed by a slight decline and a period of stability. From the early 2000s, spending increases sharply, dips briefly after 2010 and then rises again. The most recent years show a steep surge indicating a strong push in defence investment.



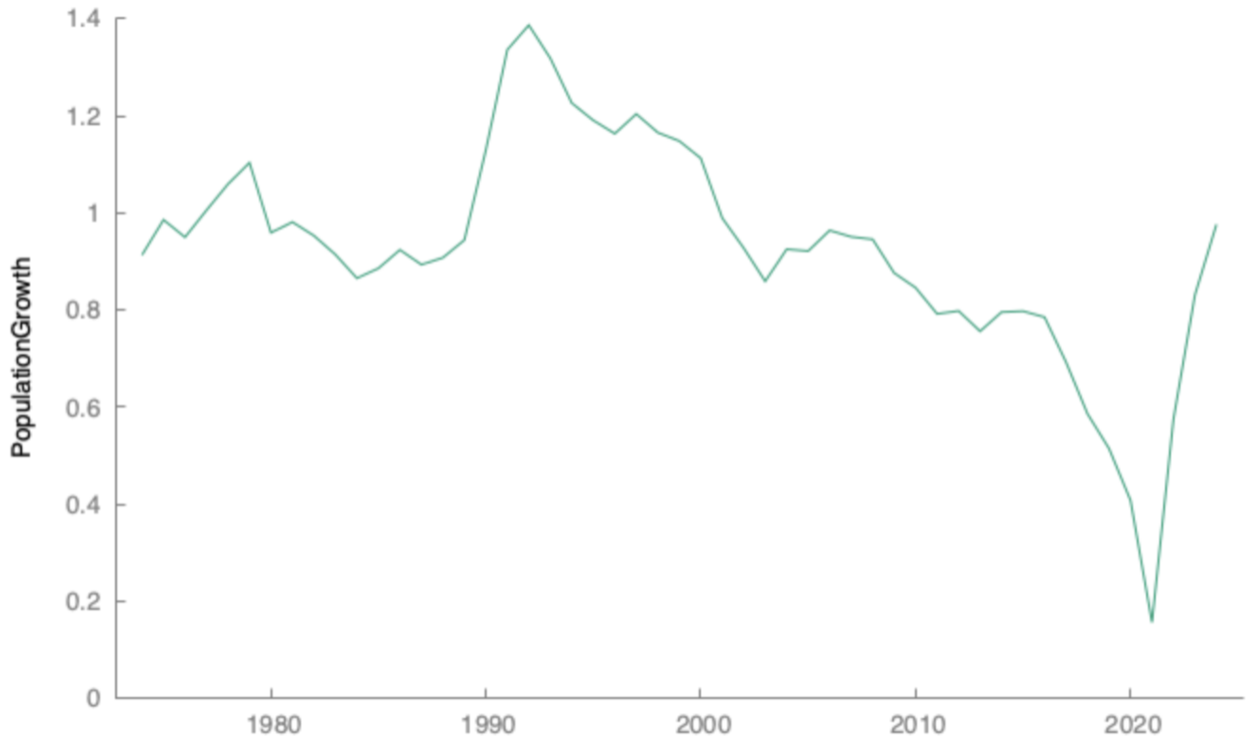
This is the time series plot of **Federal Receipts as Percentage of GDP** with time in years. **Federal Receipts as Percentage of GDP** is taken on Y-axis whereas time is taken on X-axis.

The graph shows Federal Receipts as Percentage of GDP fluctuates over time rather than following a steady trend. In the late 1970s to early 1980s, there is a general rise, reaching a peak around the early 1980s followed by a sharp drop mid-decade. Through the late 1980s and 1990s, the values move within a relatively stable range with moderate ups and downs. Around 2000, receipts peak again before falling sharply in the early 2000s. The most significant decline occurs around the 2008–2010 period, after which there is a gradual recovery. In recent years, the graph shows continued fluctuations with occasional sharp increases, indicating instability but no clear long-term upward or downward trend.



This is the time series plot of **Gross Fixed Capital Formation (Trillions of Dollars)** with time in years. **Gross Fixed Capital Formation (Trillions of Dollars)** is taken on Y-axis whereas time is taken on X-axis.

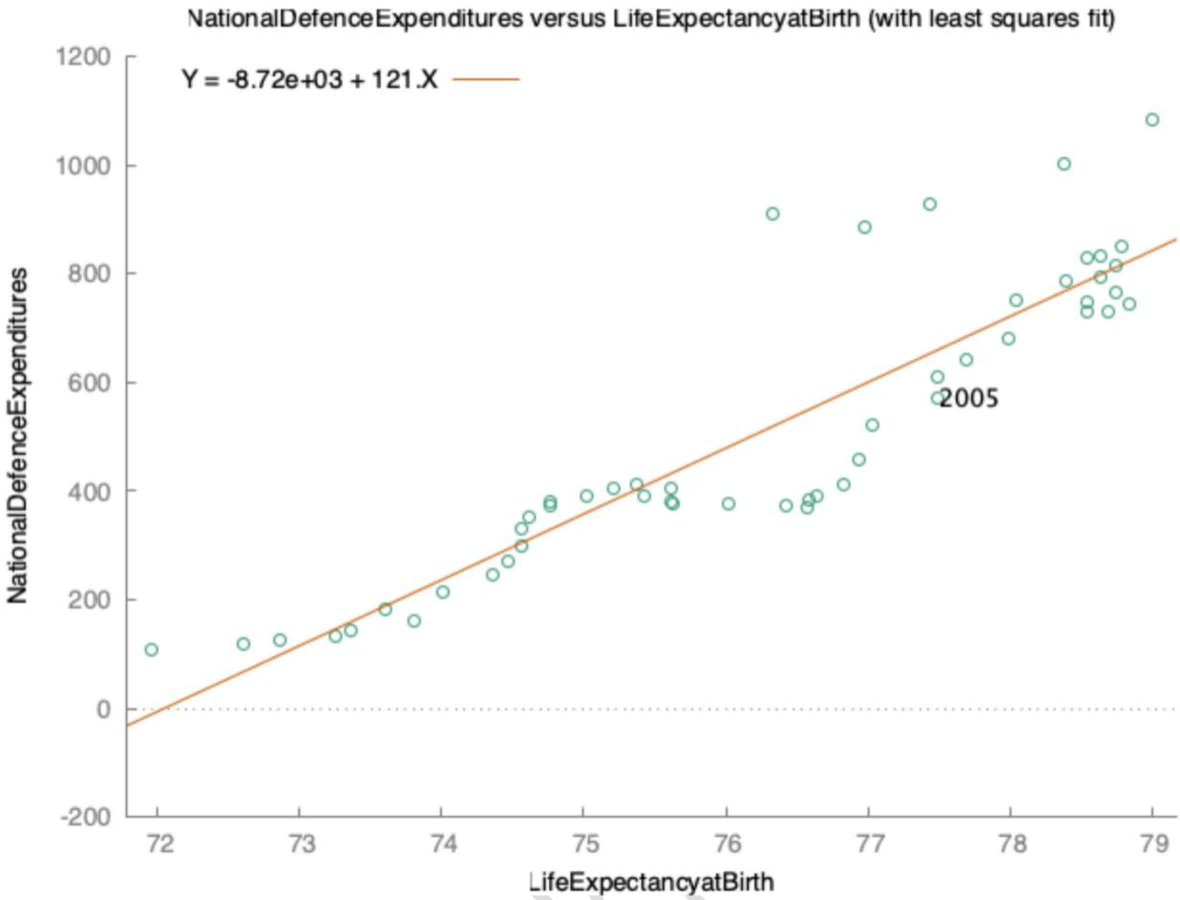
The graph shows a consistent upward trend signalling long-term growth in infrastructure and asset investment. Notable dips around 2000 and 2008 highlight the impact the global financial crisis. However, a sharp recovery after 2010 shows accelerated investment with the value quintupling over the 50-year period to reach its current peak.

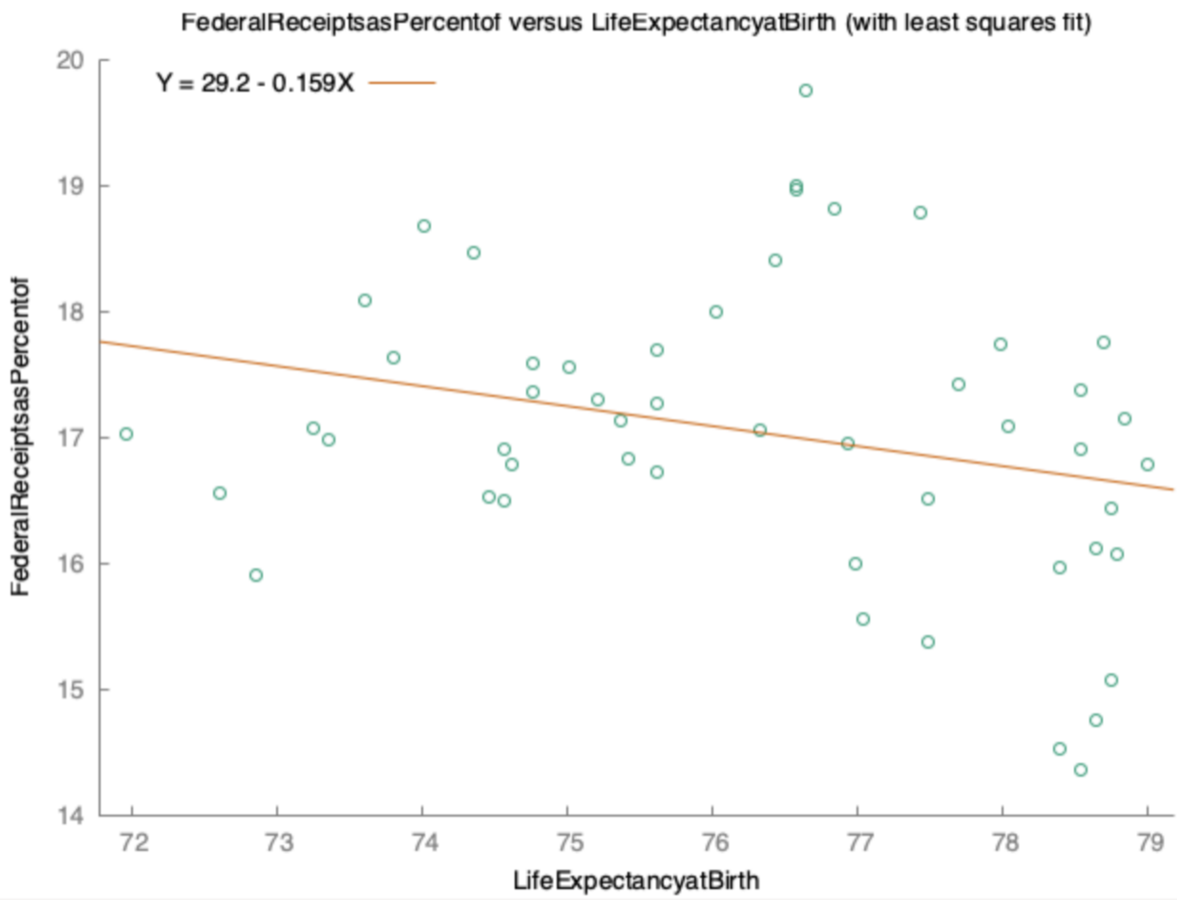


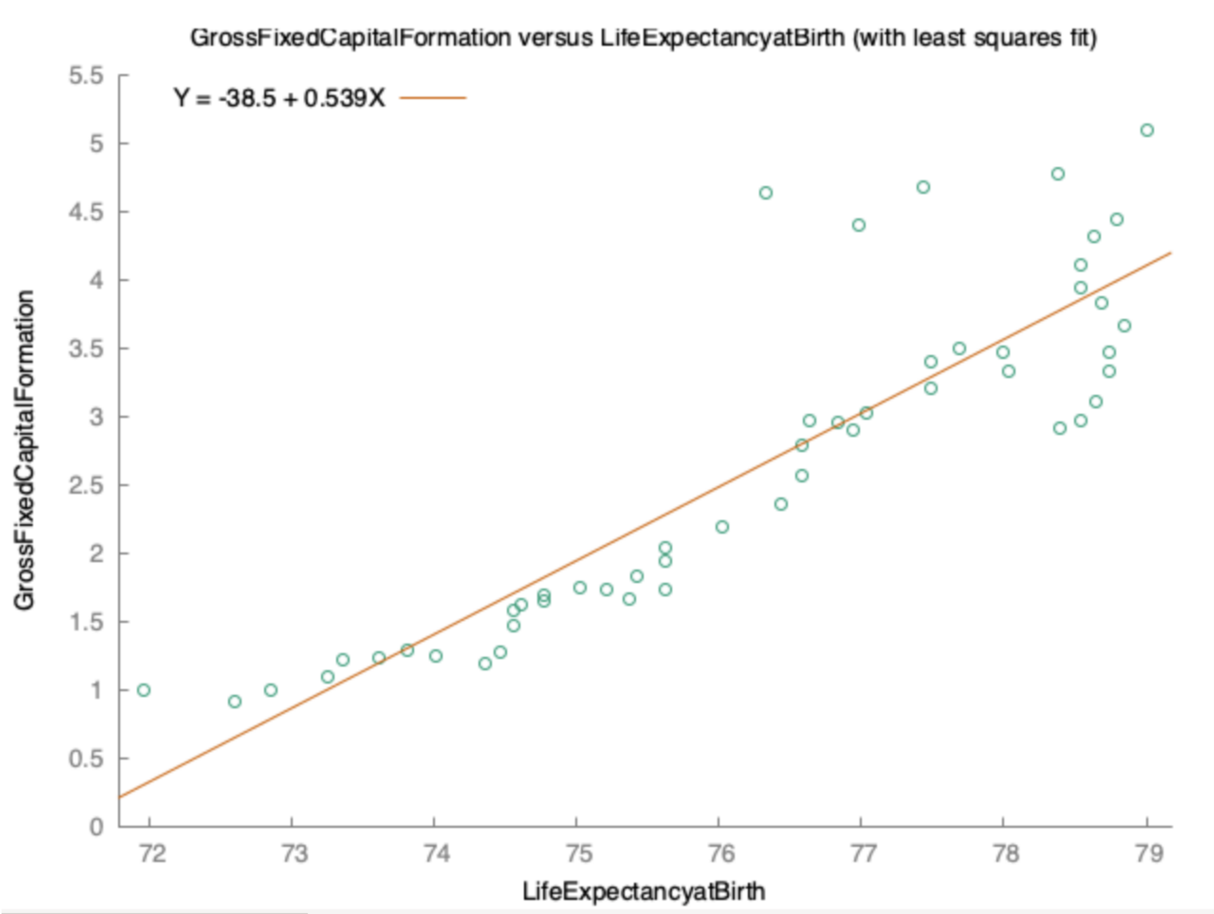
This is the time series plot of **Population Growth (Percent Change at Annual Rate)** with time in years. **Population Growth (Percent Change at Annual Rate)** is taken on Y-axis whereas time is taken on X-axis.

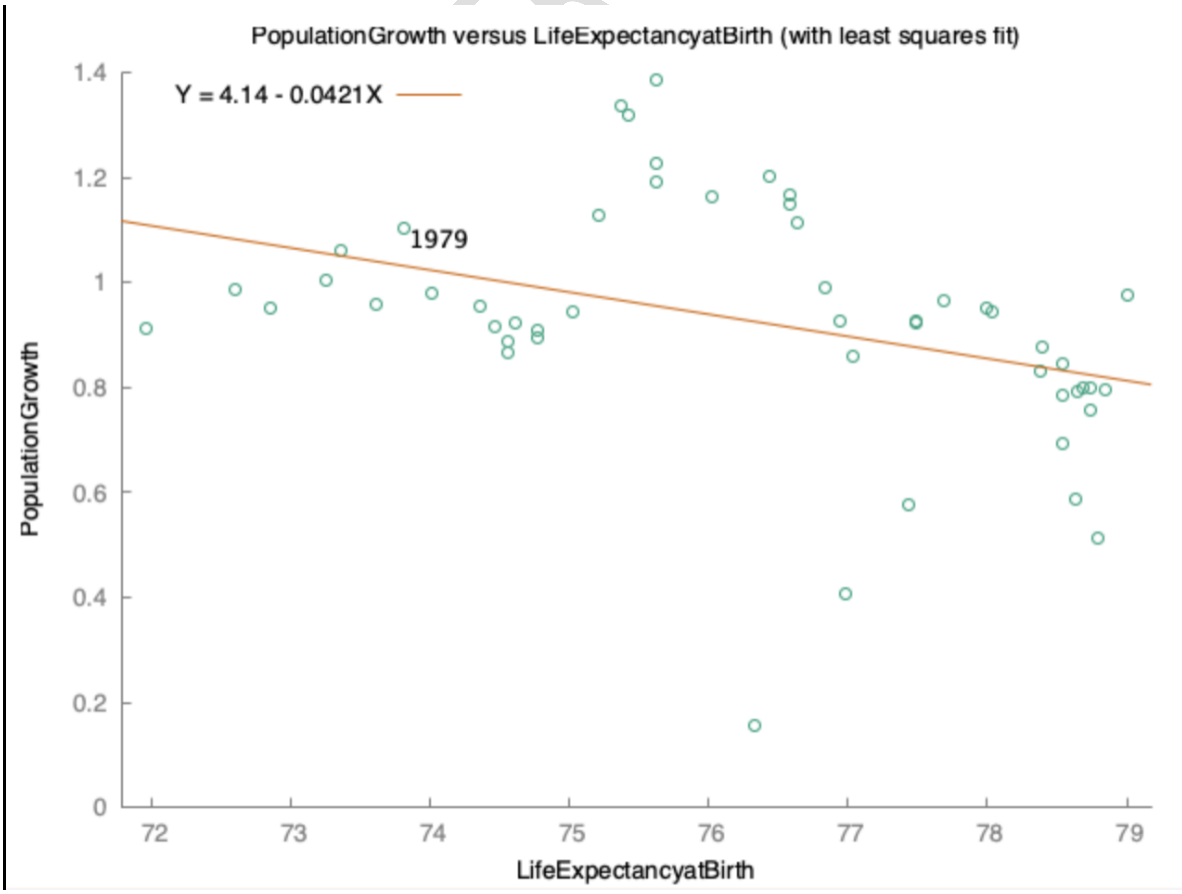
This graph illustrates the annual percentage change in Population Growth. The data reveals a generally declining trend over the decades, punctuated by a significant peak in the early 1990s where growth reached approximately 1.4%. Most notably, there is a dramatic sharp collapse around 2020 where the growth rate plummeted to near 0%, likely reflecting the demographic impact of the global pandemic. However, the tail end of the graph shows a rapid steep recovery with the growth rate surging back toward 0.8% by 2024.

SCATTER PLOTS






GrossFixedCapitalFormation versus LifeExpectancyatBirth (with least squares fit)
Y = -38.5 + 0.539X
GrossFixedCapitalFormation
LifeExpectancyatBirth
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
72
73
74
75
76
77
78
79


PopulationGrowth versus LifeExpectancyatBirth (with least squares fit)
Y = 4.14 - 0.0421X
1979
PopulationGrowth
LifeExpectancyatBirth
0
0.2
0.4
0.6
0.8
1
1.2
1.4
72
73
74
75
76
77
78
79

SETTING UP THE MODEL & HYPOTHESIS

Log_Life Expectancy at Birth = β_1 + β_2 (Log_National Defence Expenditures) -β_3 (Log_Federal Receipts as Percentage of GDP) + β_4 (Population Growth) + β_5 (Gross Fixed Capital Formation) + U_i

Setting the **Hypothesis**,

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_A : \beta_1 \neq \beta_2 \neq \beta_3 \neq \beta_4 \neq \beta_5 \neq 0$$

```
Model 1: OLS, using observations 1974-2024 (T = 51)
Dependent variable: l_LifeExpectancyatBirth

                        coefficient   std. error   t-ratio   p-value
  ------------------------------------------------------------------
  const                  4.22481     0.0621234    68.01     8.27e-48 ***
  l_NationalDefenc~      0.0253078   0.00451009    5.611    1.10e-06 ***
  l_FederalReceipt~     -0.0335601   0.0185045    -1.814    0.0763   *
  PopulationGrowth       0.0258288   0.00664747    3.886    0.0003   ***
  GrossFixedCapita~      0.0100527   0.00249981    4.021    0.0002   ***

Mean dependent var   4.334114   S.D. dependent var   0.025926
Sum squared resid    0.002820   S.E. of regression   0.007830
R-squared            0.916079   Adjusted R-squared   0.908782
F(4, 46)             125.5343   P-value(F)           3.92e-24
Log-likelihood       177.6040   Akaike criterion    -345.2081
Schwarz criterion   -335.5489   Hannan-Quinn        -341.5170
rho                  0.870109   Durbin-Watson        0.297712
```

Through the method of **Ordinary Least Squares (OLS)**,

$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \epsilon_i$

Here,

Y_i= **4.22481 + 0.0253078** X_{2i} **- 0.0335601** X_{3i} **+ 0.0258288** X_{4i} **+ 0.0100527** X_{5i} + Ui

where,

Y_i = Log_Life Expectancy at Birth

<u>Intercept</u>

β_1= 4.22481

Slopes

X_{2i}= Log_National Defence Expenditures (β_2= 0.0253078)

X_{3i}= Log_Federal Receipts as Percentage of GDP (β_3= -0.0335601)

X_{4i}= Population Growth (Percent Change at Annual Rate) (β_4= 0.0258288)

X_{5i}= Gross Fixed Capital Formation (β_5= 0.0100527)

Significance Explanations

1. **p-value**: if the p-value is less than or equal to 0.05, then we can reject the null hypothesis and the variable will be statistically significant.

2. **t-ratio:** if the absolute value of the t-ratio is more than 2, then we can reject the null hypothesis and the variable will be statistically significant.

3. **star signs** (level of significance - LOS),
- one star: rejecting the null hypothesis at 10% LOS.
- two stars: rejecting the null hypothesis at 5% & 10% LOS.
- three stars: rejecting the null hypothesis at 1%, 5% & 10% LOS.

Interpretations

Y_i= 4.22481 + 0.0253078 X_{2i} - 0.0335601 X_{3i} + 0.0258288 X_{4i} + 0.0100527 X_{5i} + Ui

- **β_1= 4.22481 (intercept value)** - When all independent variables are equal to zero, the expected value of log life expectancy is *4.22481*. Converting this into levels, the predicted life expectancy is $e^{4.22481}$ *years*. The intercept is statistically significant, as indicated by an extremely low p-value *(8.27 × 10^{-48})*, a very high t-statistic *(68.01)*, and the presence of *three stars*. This implies that the intercept is significantly different from zero, though it does not provide evidence regarding omitted variables.

- **β_2= 0.0253078 (Log_National Defence Expenditures)**- The coefficient represents an elasticity. A *1%* increase in National Defence Expenditures is associated with an approximate *0.0253078%* increase in Life Expectancy at Birth, ceteris paribus. The coefficient is positive, indicating a direct relationship between defence expenditure and life expectancy. The coefficient is statistically significant at conventional levels, as indicated by a very low p-value *(1.10 × 10^{-6})*, a t-statistic greater than 2 *(5.611)*, and the presence of *three significance stars*. Thus, the null hypothesis of no effect can be rejected with high confidence. While the coefficient is statistically significant, its magnitude suggests that the economic impact is relatively small, implying that changes in defence expenditure have a limited effect on life expectancy and hence Human Capital Development. Higher defence expenditure may be associated with higher life expectancy because it often proxies for greater state capacity, economic development and security conditions that indirectly improve health outcomes.

- **β_3= -0.0335601 (Log_Federal Receipts as Percentage of GDP)-** The coefficient represents an elasticity. A *1%* increase in Federal Receipts as a percentage of GDP is associated with an approximate *0.0336%* decrease in Life Expectancy at Birth, ceteris paribus. The coefficient is negative, indicating an inverse relationship between Federal Receipts and life expectancy. The coefficient is weakly statistically significant at the *10%* level (*p-value = 0.0763, t-statistic = -1.814*) as indicated by one significance star. Therefore, the null hypothesis of no effect can be rejected only at the *10%* level, but not at the *5%* level. Although statistically weakly significant, the magnitude of the coefficient suggests a very small economic effect implying that changes in Federal Receipts as a share of GDP have a limited impact on life expectancy and hence Human Capital Development. It is also consistent with fiscal burden and distortionary taxation literature, higher government revenue (as a share of GDP) may reflect increased tax pressures that can reduce disposable income, private consumption and private investment in health and welfare thereby potentially exerting a negative effect on long-run human capital outcomes.

- **β_4= 0.0258288 (Population Growth (Percent Change at Annual Rate))-** The coefficient represents a semi-elasticity (given the log-lin functional form), implying that a 1unit increase in population growth at annual rate in percentage terms is associated with an approximate *2.58288%* increase in Life Expectancy at Birth, ceteris paribus. The coefficient is positive, indicating a direct relationship between population growth and life expectancy. The coefficient is highly statistically significant (*p-value = 0.0003, t-statistic = 3.886, three significance stars*) implying rejection of the null hypothesis at the *1%* level. This confirms a robust relationship between population growth and life expectancy in the model. Although statistically significant, the magnitude of the coefficient suggests that the economic effect is relatively small indicating that population growth has a limited impact on life expectancy in percentage terms. Higher population growth is associated with improvements in human capital development (proxied by life expectancy), likely reflecting demographic transition effects where better health conditions simultaneously raise both survival rates and population growth.

- **β_5= 0.0100527 (Gross Fixed Capital Formation (Trillions of Dollars)**- The coefficient represents a semi-elasticity (given the log-lin functional form), implying that a *1 trillion* increase in Gross Fixed Capital Formation is associated with an approximate *1.00527%* increase in Life Expectancy at Birth, ceteris paribus. The coefficient is positive, indicating a direct relationship between capital formation and life expectancy. The coefficient is highly statistically significant (*p-value = 0.0002*, *t-statistic = 4.021, three significance stars*), implying rejection of the null hypothesis at the *1%* level. This confirms a robust relationship between Gross Fixed Capital Formation and life expectancy in the model. Although statistically significant, the magnitude of the coefficient suggests that the economic effect is relatively small indicating that increases in capital formation have a limited impact on life expectancy in percentage terms. The positive relationship is consistent with endogenous growth and development literature which argues that higher gross fixed capital formation promotes economic growth and income levels thereby indirectly improving human development outcomes and life expectancy.

- **R-squared (0.916079):** This is a measurement of the percentage of the dependent variable's overall variation that the regression model explains. A better match between the model and the data is indicated by a higher R-squared value.

- **F-statistic (125.5343) and p-value (3.92e-24)**: The regression model's overall significance is gauged by the F-statistic. Here, the regression model is very statistically significant because the p-value is very low and the F-statistic is quite high.

ANOVA

```
Analysis of Variance:

                     Sum of squares       df      Mean square

  Regression              0.0307864        4       0.00769659
  Residual               0.00282029       46      6.13107e-05
  Total                   0.0336067       50      0.000672133

  R^2 = 0.0307864 / 0.0336067 = 0.916079
  F(4, 46) = 0.00769659 / 6.13107e-05 = 125.534 [p-value 3.92e-24]
```

Mean Squares: The mean square for regression is 0.00769659, while the mean square for residuals is 6.13107e-5. The large disparity between these two values contributes to the high F-statistic and further reinforces the conclusion that the model explains a significant portion of the variation in Log_Life Expectancy at Birth.

CONFIDENCE INTERVAL

```
95% confidence intervals
t(46, 0.025) = 2.013

                               coefficient      low          high
  ---------------------------------------------------------------
  const                          4.22481      4.09976      4.34985
  l_NationalDefenceExpenditures  0.0253078    0.0162294    0.0343861
  l_FederalReceiptsasPercentof  -0.0335601   -0.0708077    0.00368758
  PopulationGrowth               0.0258288    0.0124481    0.0392094
  GrossFixedCapitalFormation     0.0100527    0.00502085   0.0150846
```

The results show that *National Defence Expenditures*, *Population Growth* and *Gross Fixed Capital Formation* significantly increase the dependent variable. All three are statistically significant at the 95% level, as their confidence intervals do not contain zero.

In contrast, *Federal Receipts as a Percentage of GDP* is not statistically significant ($p > 0.05$), meaning its effect is not clearly established. The model is highly robust, explaining 91.6% of the total variation with a highly significant overall F-test ($p < 0.001$). These findings suggest that defence and capital investment are the primary drivers in this model.

CORRELATION MATRIX

A **correlation matrix** is a square table that summarises the relationship between all pairs of variables within a dataset. It uses a statistical measure called the correlation coefficient, which ranges from *-1 to 1*.

Interpretation:

- **Values closer to 1:** Indicate a *strong positive correlation*. This means as one variable increases, the other tends to increase as well.
- **Values closer to -1:** Indicate a *strong negative correlation*. This means as one variable increases, the other tends to decrease.
- **Values closer to 0:** Indicate *a weak or no correlation*. There's little to no relationship between the changes in the two variables.

```
Correlation Coefficients, using the 51 observations 1974 - 2024
Two-tailed critical values for n = 51: 5% 0.2759, 1% 0.3575

 l_NationalDefe~ l_FederalRecei~ PopulationGrow~ GrossFixedCapi~
          1.0000         -0.2988         -0.4711          0.8969 l_NationalDefe~
                          1.0000          0.3186         -0.1767 l_FederalRecei~
                                          1.0000         -0.5975 PopulationGrow~
                                                          1.0000 GrossFixedCapi~
```

Based on the correlation matrix provided, the calculated correlation coefficients are as follows:

- Correlation coefficient between *Log_National Defence Expenditures* and *Log_Federal Receipts as Percentage of GDP* is *−0.2988 — Weak Negative Correlation*.

- Correlation coefficient between *Log_National Defence Expenditures* and *Population Growth* is *−0.4711 — Moderate Negative Correlation*.

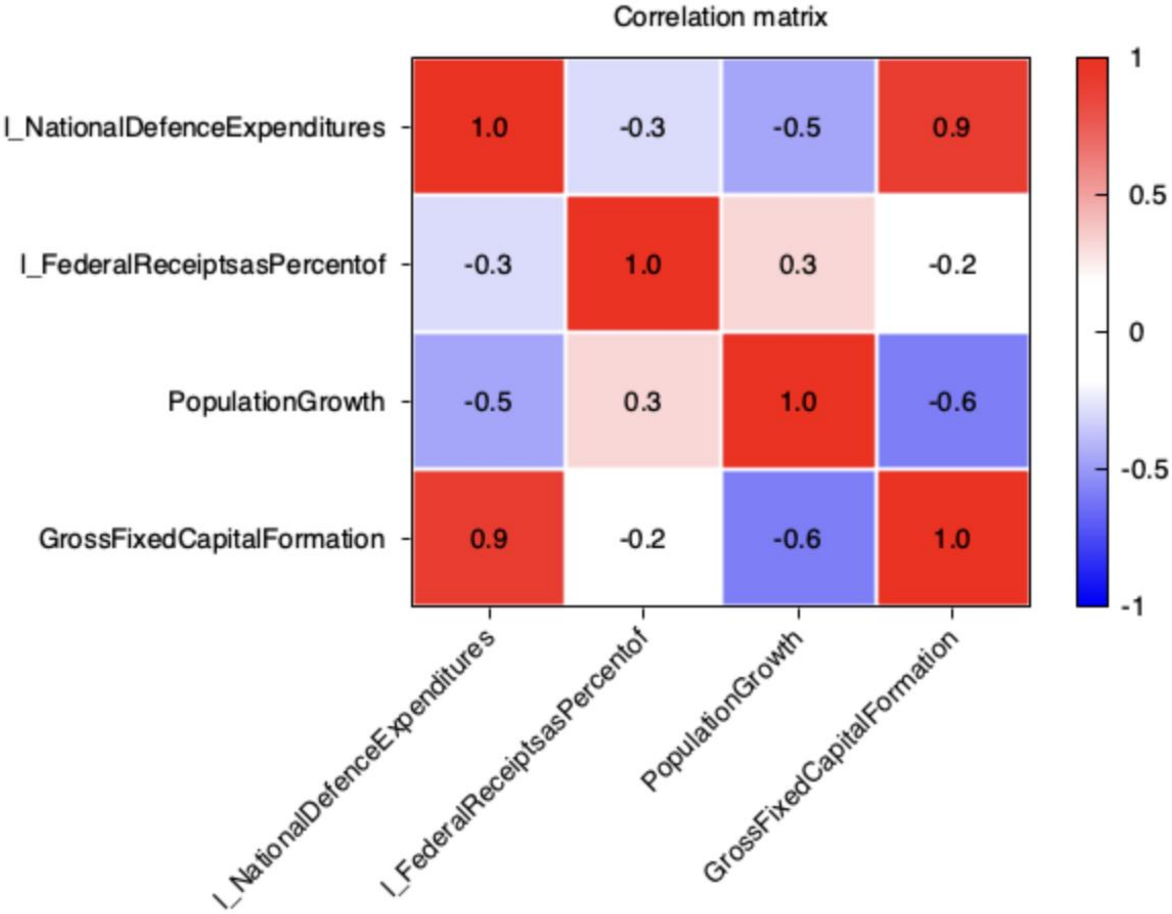
- Correlation coefficient between *Log_National Defence Expenditures* and *Gross Fixed Capital* is *0.8969 — Strong Positive Correlation*.

- Correlation coefficient between *Log_Federal Receipts as Percentage of GDP* and *Population Growth* is *0.3186 — Weak Positive Correlation*.

- Correlation coefficient between *Log_Federal Receipts as Percentage of GDP* and *Gross Fixed Capital* is *−0.1767 — Very Weak Negative*.

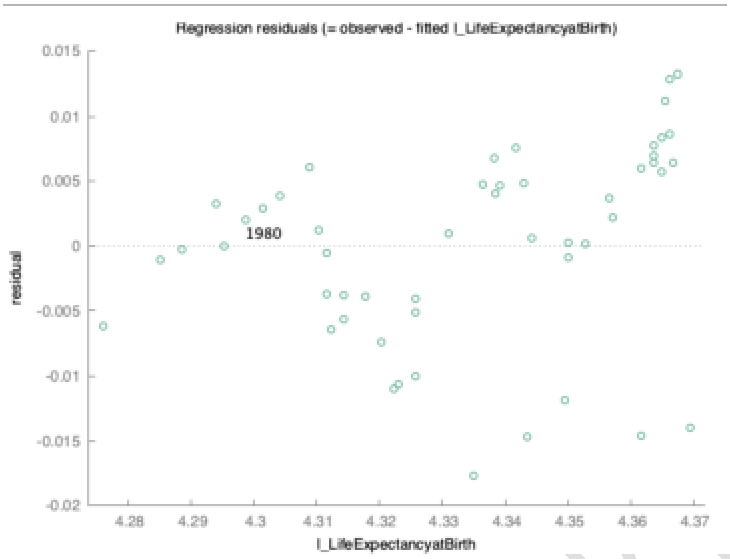
- Correlation coefficient between *Population Growth* and *Gross Fixed Capital* is *−0.5975 — Moderate Negative Correlation*.

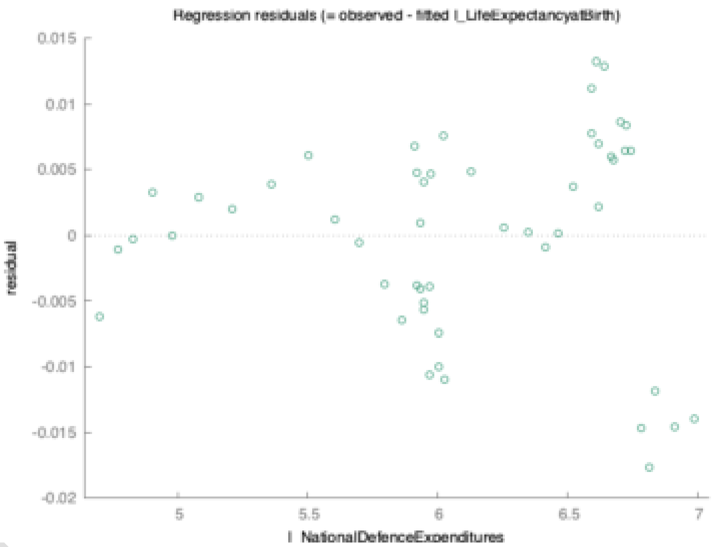
HeatMap of the Correlation Matrix:

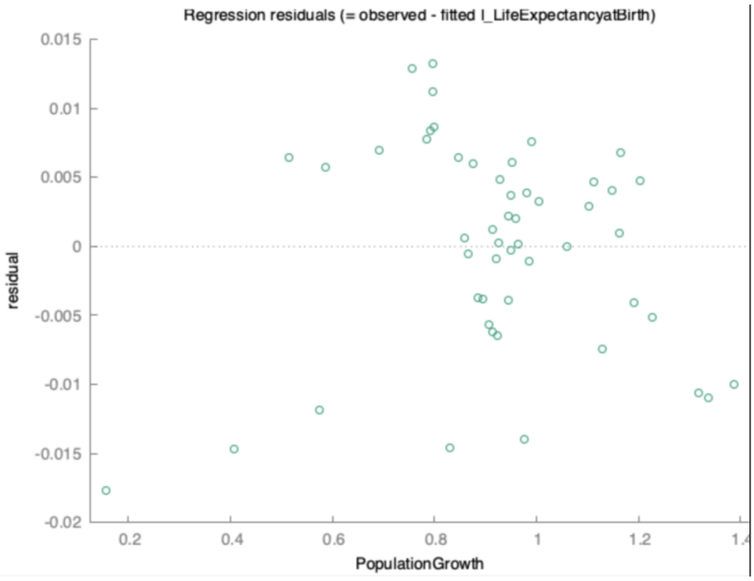


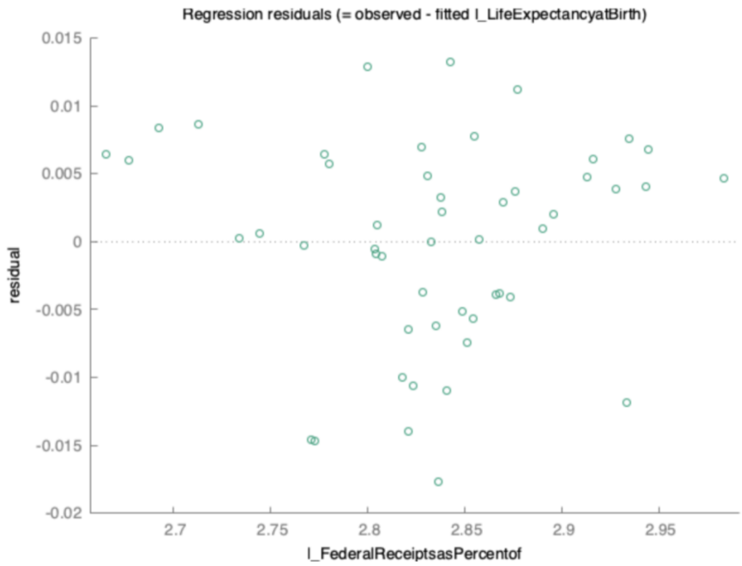
RESIDUAL ANALYSIS

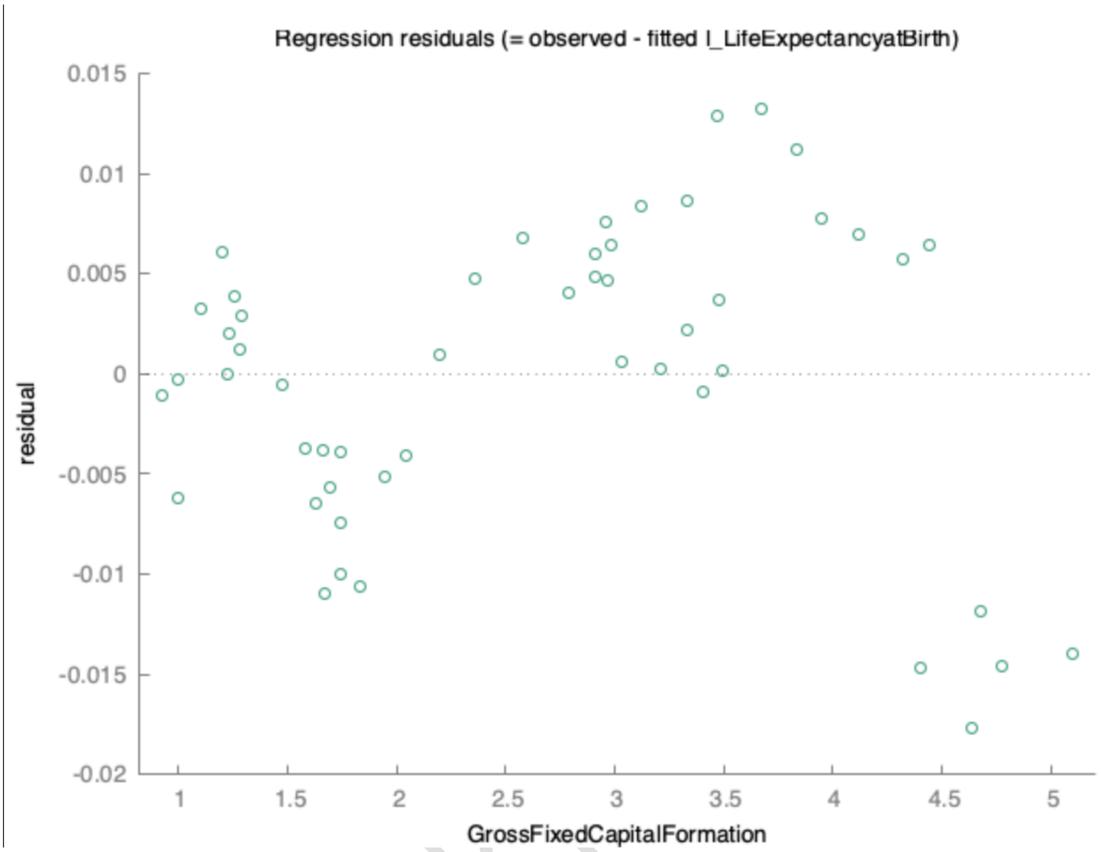
A **residual analysis** is done to check the model's fit to the data. A residual is a *difference between the actual value and the predicted value* of the dependent variable. By analysing the residuals, we can verify if the model fits the data well and identify any pattern related to outliers in the data. As we know, there is no pattern for the residuals and thus we can check it through the residual plots.









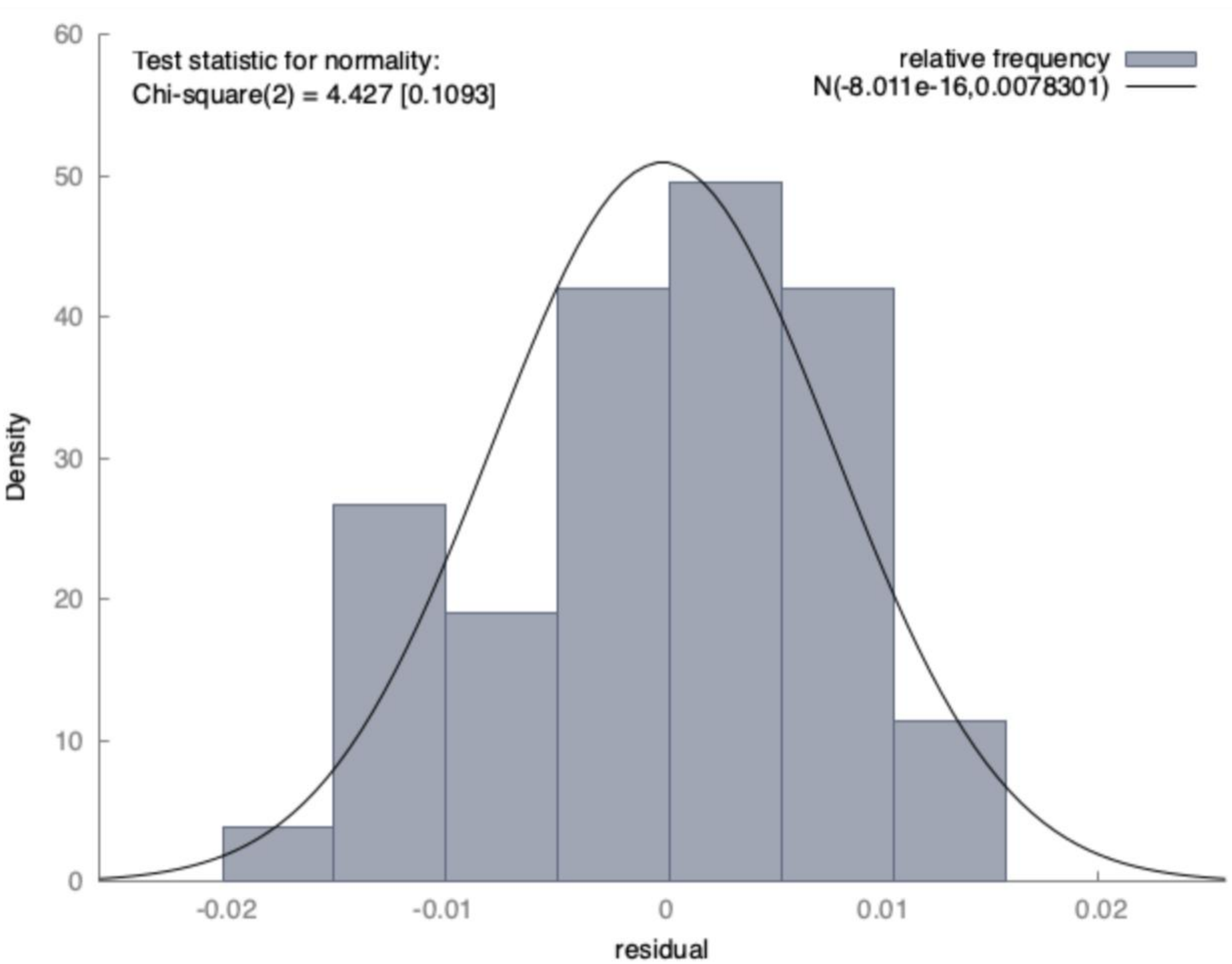


Upon looking at the graphs, it can be clearly stated that the **residuals do not exhibit a fixed pattern** or do not follow any kind of trend. This *proves the fact that errors or residuals have no pattern.*

NORMALITY OF RESIDUALS

H_0: $Residuals\ follow\ a\ normal\ distribution$

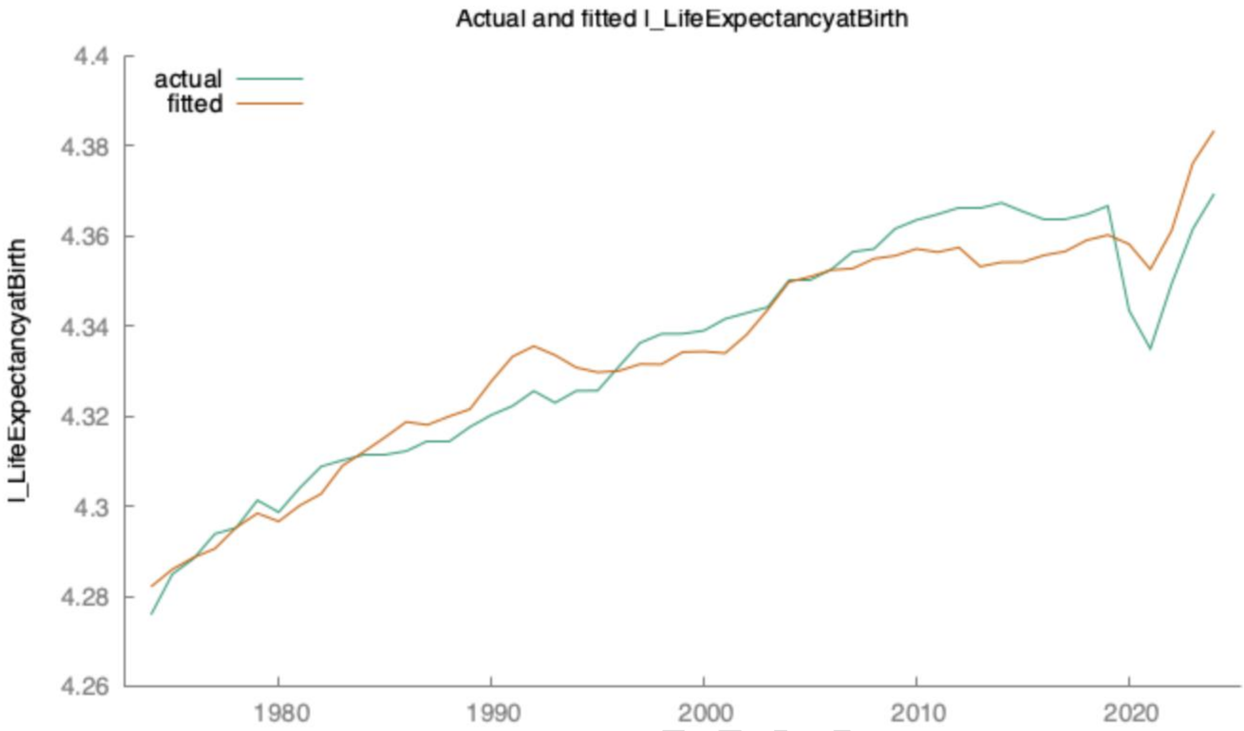
H_1: $Residuals\ do\ not\ follow\ a\ normal\ distribution$



The curve for the test of normality, thus shows that the **residuals do follow a normal distribution**.
This hypothesis test is conducted with the help of chi-square test in order to know the normality of the residuals.
Test-statistic: Chi-square(2) = *4.427* & p-value: *0.1093*

Interpretation: The *p-value is greater than the test statistic*. Thus, we fail to reject the null hypothesis, which in turn means the **residuals follow a normal distribution**.

FITTED vs ACTUAL



In regression analysis, the ***actual values are the observed outcomes*** *of the dependent variable while the* ***fitted values are the predictions*** *made by the regression model based on a given set of independent variables. Assessing the disparity between these two values serves as a means to gauge the model's accuracy.*

A well-fitting model typically yields actual and fitted values that closely align, with small and random residuals. Conversely, a poorly fitting model results in larger discrepancies between actual and fitted values, accompanied by discernible patterns in the residuals, indicating the model's failure to accurately capture the interrelationships among variables. In this specific model, the proximity between actual and fitted values suggests a robust correlation, indicating a strong alignment between the model's predictions and observed outcomes.

HETEROSKEDASTICITY - WHITE'S TEST

H_0: $Heteroskedasticity\ not\ present$ H_1: $Heteroskedasticity\ is\ present$

```
White's test for heteroskedasticity
OLS, using observations 1974-2024 (T = 51)
Dependent variable: uhat^2

                         coefficient   std. error   t-ratio  p-value
  -----------------------------------------------------------------
  const                  -0.00633200   0.0227457    -0.2784  0.7823
  l_NationalDefenc~      -0.000935882  0.00273247   -0.3425  0.7340
  l_FederalReceipt~       0.00492869   0.0130529     0.3776  0.7080
  PopulationGrowth        0.00232276   0.00370312    0.6272  0.5345
  GrossFixedCapita~       0.000281724  0.00158490    0.1778  0.8599
  sq_l_NationalDef~       9.69124e-05  0.000151264   0.6407  0.5258
  X2_X3                   8.58117e-05  0.000721297   0.1190  0.9060
  X2_X4                  -0.000281059  0.000311980  -0.9009  0.3736
  X2_X5                  -2.99773e-05  0.000172434  -0.1738  0.8630
  sq_l_FederalRece~      -0.000719681  0.00186306   -0.3863  0.7016
  X3_X4                  -0.000801745  0.00100100   -0.8009  0.4284
  X3_X5                  -0.000151779  0.000389391  -0.3898  0.6990
  sq_PopulationGro~       0.000622810  0.000197291   3.157   0.0032  ***
  X4_X5                   0.000204978  0.000136210   1.505   0.1411
  sq_GrossFixedCap~       2.08883e-05  4.28193e-05   0.4878  0.6286

  Unadjusted R-squared = 0.744828

Test statistic: TR^2 = 37.986229,
with p-value = P(Chi-square(14) > 37.986229) = 0.000522
```

```
White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 37.9862
  with p-value = P(Chi-square(14) > 37.9862) = 0.000522183
```

Interpretation: The *p-value is very small*, therefore we have enough evidence to **reject the null hypothesis**. Therefore, **heteroskedasticity is present in the model**.

MULTICOLLINEARITY

```
Variance Inflation Factors
Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

l_NationalDefenceExpenditures    6.257
 l_FederalReceiptsasPercentof    1.316
             PopulationGrowth    1.839
   GrossFixedCapitalFormation    7.334

VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient
between variable j and the other independent variables

Belsley-Kuh-Welsch collinearity diagnostics:

  variance proportions

    lambda      cond     const l_Nation~ l_Federa~ Populati~ GrossFix~
     4.797     1.000     0.000     0.000     0.000     0.001     0.001
     0.183     5.120     0.000     0.000     0.000     0.050     0.081
     0.018    16.508     0.003     0.004     0.004     0.842     0.169
     0.002    52.360     0.005     0.635     0.059     0.003     0.485
     0.000   163.528     0.992     0.361     0.938     0.104     0.264

  lambda = eigenvalues of inverse covariance matrix (smallest is 0.000179403)
  cond   = condition index
  note: variance proportions columns sum to 1.0

According to BKW, cond >= 30 indicates "strong" near linear dependence,
and cond between 10 and 30 "moderately strong".  Parameter estimates whose
variance is mostly associated with problematic cond values may themselves
be considered problematic.

Count of condition indices >= 30: 2
Variance proportions >= 0.5 associated with cond >= 30:

     const l_Nation~ l_Federa~ GrossFix~
     0.997     0.996     0.996     0.749

Count of condition indices >= 10: 3
Variance proportions >= 0.5 associated with cond >= 10:

     const l_Nation~ l_Federa~ Populati~ GrossFix~
     1.000     1.000     1.000     0.948     0.918
```

Interpretation: All the variables have **Variance Inflation Factors** less than *10* and hence, there is **no multicollinearity** among the independent variables in the model. It does not raise any concern.

AUTOCORRELATION

Autocorrelation is a representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. As our data is measured annually, the lagged intervals will be 1. Two tests will be conducted for autocorrelation.

Breusch-Godfrey Test

H_0: no
$autocorrelation$ H_A:
$not\ H_0$

```
Breusch-Godfrey test for first-order autocorrelation
OLS, using observations 1974-2024 (T = 51)
Dependent variable: uhat

                       coefficient   std. error  t-ratio   p-value
  ---------------------------------------------------------------
  const                -0.0289420    0.0324502   -0.8919   0.3772
  l_NationalDefenc~     0.000635577  0.00234901   0.2706   0.7880
  l_FederalReceipt~     0.00713595   0.00965613   0.7390   0.4637
  PopulationGrowth      0.00523409   0.00349282   1.499    0.1410
  GrossFixedCapita~    -9.44996e-05  0.00130163  -0.07260  0.9424
  uhat_1                0.907619     0.0812862   11.17     1.47e-14 ***

  Unadjusted R-squared = 0.734784

Test statistic: LMF = 124.673282,
with p-value = P(F(1,45) > 124.673) = 1.47e-14

Alternative statistic: TR^2 = 37.474005,
with p-value = P(Chi-square(1) > 37.474) = 9.26e-10

Ljung-Box Q' = 35.4314,
with p-value = P(Chi-square(1) > 35.4314) = 2.64e-09
```

Interpretation: The *p-value is very small*, therefore we have enough evidence to **reject the null hypothesis**. Therefore, **autocorrelation is present in the model**.

<u>Durbin-Watson Test</u>

H_0: no
$autocorrelation$ H_A :
not H_0

```
Durbin–Watson statistic = 0.297712

H1: positive autocorrelation
   p–value = 9.15878e–13
H1: negative autocorrelation
   p–value = 1
```


Interpretation: The **Durbin-Watson statistic** for this model is *0.2977* which is substantially **lower** than the *ideal value of 2.0*. Because the statistic is close to zero and the associated p-value ($p < 0.05$) is **highly significant**, we **reject the null hypothesis of no autocorrelation**. We conclude that there is strong evidence of **positive first-order autocorrelation** in the residuals meaning the error terms are not independent over time.

CHOW TEST

The **Chow Test** is used to test whether the structural breaks are statistically significant or not.

H_0: $no\ structural$
$break\ H_A$: $not\ H_0$

Structural Break in Year 2001

The **year 2001** is introduced as a potential structural break point in the model to account for the **September 11 terrorist attacks**, which marked a major geopolitical and economic turning point in the United States. The event led to a significant shift in national priorities, particularly with respect to defence expenditure, national security policies, and government spending patterns. These changes likely influenced broader macroeconomic conditions, including investment decisions and fiscal allocations, which may in turn affect human capital development proxied by life expectancy.

```
Augmented regression for Chow test
OLS, using observations 1974-2024 (T = 51)
Dependent variable: l_LifeExpectancyatBirth

                      coefficient   std. error   t-ratio   p-value
  -----------------------------------------------------------------
  const                4.09258      0.0979975    41.76     3.17e-35 ***
  l_NationalDefenc~    0.0181317    0.00398720    4.547    4.74e-05 ***
  l_FederalReceipt~    0.0279833    0.0321621     0.8701   0.3893
  PopulationGrowth     0.0155358    0.00878585    1.768    0.0845   *
  GrossFixedCapita~    0.0135687    0.00450113    3.015    0.0044   ***
  splitdum             0.0384238    0.144234      0.2664   0.7913
  sd_l_NationalDef~    0.0173564    0.0105034     1.652    0.1061
  sd_l_FederalRece~   -0.0355472    0.0395686    -0.8984   0.3742
  sd_PopulationGro~    0.0109256    0.0114422     0.9549   0.3452
  sd_GrossFixedCap~   -0.0155971    0.00583611   -2.673    0.0108   **

Mean dependent var   4.334114   S.D. dependent var   0.025926
Sum squared resid    0.001358   S.E. of regression   0.005755
R-squared            0.959588   Adjusted R-squared   0.950717
F(9, 41)             108.1717   P-value(F)           1.14e-25
Log-likelihood       196.2379   Akaike criterion    -372.4758
Schwarz criterion   -353.1575   Hannan-Quinn        -365.0937
rho                  0.683898   Durbin-Watson        0.578344

Chow test for structural break at observation 2001
  F(5, 41) = 8.82825 with p-value 0.0000
```

```
Chow test for structural break at observation 2001 -
  Null hypothesis: no structural break
  Test statistic: F(5, 41) = 8.82825                        \
  with p-value = P(F(5, 41) > 8.82825) = 9.47418e-06
```

Interpretation: The p-value of the test is less than 0.05 i.e. *9.47418e-06*, which makes the 2001 **September 11 terrorist attacks** a *statistically significant structural break* i.e. *2001*.

Structural Break in Year 2008

The year **2008** is introduced as a potential structural break point in the model to account for the **Global Financial Crisis**, which marked one of the most severe economic downturns in modern US history. The crisis led to a collapse in financial markets, a sharp contraction in credit availability, rising unemployment, and significant disruptions in investment, fiscal balances, and overall economic stability. Given these widespread macroeconomic changes, it is plausible that the relationships between the explanatory variables and life expectancy (as a proxy for human capital development) may have experienced a structural shift post-2008. Therefore, a Chow test is employed to examine whether the regression coefficients remain stable before and after the crisis period.

```
Augmented regression for Chow test
OLS, using observations 1974-2024 (T = 51)
Dependent variable: l_LifeExpectancyatBirth

                      coefficient  std. error  t-ratio   p-value
  ---------------------------------------------------------------
  const                4.14782     0.0383899   108.0     5.58e-52 ***
  l_NationalDefenc~    0.0172520   0.00252116    6.843   2.75e-08 ***
  l_FederalReceipt~    0.0101016   0.0127278     0.7937  0.4320
  PopulationGrowth     0.0129967   0.00502941    2.584   0.0134   **
  GrossFixedCapita~    0.0157600   0.00156958   10.04    1.30e-12 ***
  splitdum             0.755185    0.143837      5.250   5.00e-06 ***
  sd_l_NationalDef~   -0.0786612   0.0157458    -4.996   1.14e-05 ***
  sd_l_FederalRece~   -0.0850016   0.0249883    -3.402   0.0015   ***
  sd_PopulationGro~    0.0282886   0.00740713    3.819   0.0004   ***
  sd_GrossFixedCap~   -0.00308312  0.00374089   -0.8242  0.4146

Mean dependent var   4.334114   S.D. dependent var   0.025926
Sum squared resid    0.000619   S.E. of regression   0.003887
R-squared            0.981569   Adjusted R-squared   0.977523
F(9, 41)             242.6143   P-value(F)           1.26e-32
Log-likelihood       216.2581   Akaike criterion    -412.5161
Schwarz criterion   -393.1979   Hannan-Quinn        -405.1340
rho                  0.234276   Durbin-Watson        1.394408

Chow test for structural break at observation 2008
  F(5, 41) = 29.1368 with p-value 0.0000
```

```
Chow test for structural break at observation 2008 -
  Null hypothesis: no structural break
  Test statistic: F(5, 41) = 29.1368
  with p-value = P(F(5, 41) > 29.1368) = 1.71324e-12
```

Interpretation: The p-value of the test is less than 0.05 i.e. *1.71324e-12*, which makes the **2008 Global Financial Crisis** a *statistically significant structural break* i.e. *2008.*

Structural Break in Year 2020

The year 2020 marks a potential structural break in the United States due to the outbreak of the COVID-19 pandemic, which caused an unprecedented health and economic shock. The pandemic led to a sharp decline in life expectancy, significant disruptions in economic activity, and major changes in fiscal policy, healthcare expenditure, and labour market conditions. Given these large-scale and sudden shifts, 2020 is a suitable break point to examine whether the relationships between macroeconomic variables and life expectancy changed significantly in the post-pandemic period.

```
Augmented regression for Chow test
OLS, using observations 1974-2024 (T = 51)
Dependent variable: l_LifeExpectancyatBirth

                      coefficient   std. error    t-ratio    p-value
  ------------------------------------------------------------------
  const                 4.20810     0.0372593   112.9       9.12e-53 ***
  l_NationalDefenc~     0.0219245   0.00263117    8.333     2.33e-10 ***
  l_FederalReceipt~    -0.0187793   0.0114205    -1.644     0.1078
  PopulationGrowth      0.0120844   0.00485226    2.490     0.0169   **
  GrossFixedCapita~     0.0141576   0.00158888    8.910     3.87e-11 ***
  splitdum              1.06355     8.89474       0.1196    0.9054
  sd_l_NationalDef~    -0.172366    1.42334      -0.1211    0.9042
  sd_l_FederalRece~    -0.0255437   0.286941     -0.08902   0.9295
  sd_PopulationGro~     0.0359506   0.108023      0.3328    0.7410
  sd_GrossFixedCap~     0.0303695   0.343648      0.08837   0.9300

Mean dependent var   4.334114   S.D. dependent var   0.025926
Sum squared resid    0.000774   S.E. of regression   0.004345
R-squared            0.976964   Adjusted R-squared   0.971907
F(9, 41)             193.1988   P-value(F)           1.20e-30
Log-likelihood       210.5703   Akaike criterion    -401.1406
Schwarz criterion   -381.8223   Hannan-Quinn        -393.7585
rho                  0.726559   Durbin-Watson        0.482872

Chow test for structural break at observation 2020
  F(5, 41) = 21.6722 with p-value 0.0000
```

```
Chow test for structural break at observation 2020 -
  Null hypothesis: no structural break
  Test statistic: F(5, 41) = 21.6722
  with p-value = P(F(5, 41) > 21.6722) = 1.4964e-10
```

Interpretation: The p-value of the test is less than 0.05 i.e. *1.4964e-10*, which makes the **COVID-19 pandemic** a *statistically significant structural break* i.e. 2020

RESET SPECIFICATION TEST

The **RESET Specification Test** is used to test whether the model needs a new functional form or not and if it suffers from any omitted variables or not.
The null hypothesis indicates that the functional form being used is adequate and there are no omitted variables.

$H_0: Specification\ is$

$adequate\ H_A: not\ H_0$

```
Auxiliary regression for RESET specification test
OLS, using observations 1974-2024 (T = 51)
Dependent variable: l_LifeExpectancyatBirth

                       coefficient   std. error    t-ratio   p-value
  ------------------------------------------------------------------
  const                76.8191      44.8825        1.712    0.0939   *
  l_NationalDefenc~     0.908810     0.546256      1.664    0.1031
  l_FederalReceipt~   -1.23381       0.742292     -1.662    0.1034
  PopulationGrowth      0.935419     0.562406      1.663    0.1032
  GrossFixedCapita~     0.367409     0.220954      1.663    0.1033
  yhat^2               -0.00250266   0.00154731   -1.617    0.1128

Null hypothesis: specification is adequate
Test statistic: F = 2.616080,
with p-value = P(F(1,45) > 2.61608) = 0.113
```

```
RESET test for specification (squares only) -
  Null hypothesis: specification is adequate
  Test statistic: F(1, 45) = 2.61608
  with p-value = P(F(1, 45) > 2.61608) = 0.112775
```

Interpretation: The p-value of the test is more than 0.05 i.e. *0.113*. Thus, we **fail to reject the null hypothesis**, meaning, **the functional form used is correct** and our *model does not suffer from any omitted variables*.

LENIER RESTRICTIONS TEST

The Linear Restrictions Test is used to examine specific hypotheses about the relationships between regression coefficients. Unlike the overall F-test, which checks whether all explanatory variables jointly influence the dependent variable, this test allows us to impose and evaluate particular restrictions based on economic reasoning. It helps determine whether certain variables have equal effects, whether their combined impact follows a theoretical expectation, or whether a specific variable can be excluded from the model. The test is conducted using the F-statistic framework, and the results provide deeper insights into the structure and validity of the estimated regression model.

To further examine specific relationships between the explanatory variables, linear restrictions were tested using the F-test in Gretl.

- To compare the impact of defence expenditure and gross fixed capital formation:

$$H_0: \beta_2 = \beta_5$$
$$H_A: \beta_2 \neq \beta_5$$

```
Restriction:
 b[l_NationalDefenceExpenditures] - b[GrossFixedCapitalFormation] = 0

Test statistic: F(1, 46) = 4.99345, with p-value = 0.0303428

Restricted estimates:

                                 coefficient  std. error   t-ratio  p-value
  ----------------------------------------------------------------------
  const                           4.31152     0.0505333    85.32    3.42e-53 ***
  l_NationalDefenceExpenditures   0.0153739   0.000792314  19.40    4.48e-24 ***
  l_FederalReceiptsasPercentof   -0.0496795   0.0177498    -2.799   0.0074   ***
  PopulationGrowth                0.0315515   0.00638966    4.938   1.04e-05 ***
  GrossFixedCapitalFormation      0.0153739   0.000792314  19.40    4.48e-24 ***

  Standard error of the regression = 0.00815598
```

INTERPRETATION: An F-test was performed to test the equality of the coefficients for National Defence Expenditures and Gross Fixed Capital Formation. With an *F-statistic of 4.99* and a *p-value of 0.03*, **we reject the null hypothesis of equality** at the 5% level. This indicates that these two factors do not have the same marginal effect on the dependent variable

- To examine fiscal trade-offs:

$$H_0: \beta_2 + \beta_5 = 0$$
$$H_A: \beta_2 + \beta_5 \neq 0$$

```
Restriction:
 b[l_NationalDefenceExpenditures] + b[l_FederalReceiptsasPercentof] = 0

Test statistic: F(1, 46) = 0.159, with p-value = 0.691924

Restricted estimates:

                                  coefficient  std. error  t-ratio  p-value
  ---------------------------------------------------------------------------
  const                            4.20038     0.0102754   408.8    4.12e-85 ***
  l_NationalDefenceExpenditures    0.0263311   0.00367540    7.164  4.61e-09 ***
  l_FederalReceiptsasPercentof    -0.0263311   0.00367540   -7.164  4.61e-09 ***
  PopulationGrowth                 0.0248204   0.00609244    4.074  0.0002   ***
  GrossFixedCapitalFormation       0.00954248  0.00212819    4.484  4.70e-05 ***

  Standard error of the regression = 0.00775974
```

INTERPRETATION: To test whether the combined impact of National Defence Expenditures and Federal Receipts is net zero, a linear restriction test was conducted. The resulting F-statistic (0.159) with a p-value of 0.69 is well above the 5% significance level. Therefore, we fail to reject the null hypothesis. This suggests that the data does not provide sufficient evidence against the imposed restriction, implying that the assumption of a net-zero combined effect cannot be rejected.

CONCLUSION

In this Econometric Study, several key insights emerged regarding the factors influencing the **Life Expectancy at Birth (Total for the United States; no. of years) (Proxy for Human Capital Development)** in the USA. The study took a **Time Series** approach, focusing on the Annual data from **1974 to 2024** to examine and dissect the **relationship between** the **Life Expectancy at Birth (Proxy for Human Capital Development)** and four independent variables - **National Defence Consumption Expenditures and Gross Investment (Billions of Dollars), (2) Government Total Revenue (Percent of Gross Domestic Product), (3) Population Growth Rate (Percent Change at Annual Rate), (4) Gross Fixed Capital Formation (trillions of Dollars).**

The Regression is as follows

Log_Life Expectancy at Birth = β_1 + β_2 (Log_National Defence Expenditures) -β_3 (Log_Federal Receipts as Percentage of GDP) + β_4 (Population Growth) + β_5 (Gross Fixed Capital Formation) + U_i

Y_i= **4.22481 + 0.0253078** X_{2i} **- 0.0335601** X_{3i} **+ 0.0258288** X_{4i} **+ 0.0100527** X_{5i} + Ui

The analysis revealed several significant findings. *Log National Defence Expenditures* and *Gross Fixed Capital Formation* exhibited a s*trong positive correlation* with *Life Expectancy at Birth*, indicating that higher government spending capacity and capital investment are associated with improved life expectancy. On the other hand, *Population Growth* showed a *negative correlation* with *Life Expectancy*, suggesting that higher population growth may put pressure on resources and reduce overall health outcomes. *Log Federal Receipts as a Percentage of GDP* also displayed a *weak negative correlatio*n with *Life Expectancy*, although the relationship is limited. While the *positive association* of *defence expenditure* with l*ife expectancy* may seem unexpected, it likely reflects overall economic development rather than a direct causal effect.

The model's *constant term* is statistically *significant* indicating that the baseline level of Life Expectancy is well captured even in the absence of explanatory variables. Furthermore, the overall specification of the model appears appropriate as supported by diagnostic testing such as the RESET test which suggests that the chosen functional form is adequate and that there is no strong evidence of omitted variable bias. This implies that the variables included in the model are sufficient to explain variations in Life Expectancy.

The model's *high R-squared value* of *0.916079* indicates that the independent variables explain approximately *91.6079%* of the variation in Life Expectancy at Birth, highlighting the model's strong explanatory power. Furthermore, the *F-statistic* of *125.5343* and its associated *p-value* of *3.92e-24* confirm the overall significance of the model.

I conducted various tests for the regression model. Firstly, I examined the *Normality of Residuals*, which indicated that the *residuals were distributed normally*. Given that my study follows a Time Series approach, *heteroskedasticity* and *autocorrelation* is *expected to occur.* However, my model did *not show any sign of multicollinearity*, which was confirmed through the *Variance Inflation Factor (VIF) test*. Additionally, I conducted *three Chow tests* with the structural breaks identified as *2001 (September attacks)*, *2008 (the Global Financial Crisis)* and *2020 (Covid-19)*. All of these structural breaks were found to be *statistically significant*.

Life expectancy is influenced by a wide range of factors that are not always feasible to fully capture or predict. However, this study focuses on key macroeconomic variables that significantly shape human capital development in the United States. In particular, the analysis highlights the relationship between defence expenditure and life expectancy as a proxy for human capital development. The findings suggest that higher defence expenditure is positively associated with life expectancy which may reflect broader economic strength, technological advancement and institutional capacity that indirectly support public health systems and living standards.

At the same time, this relationship is not necessarily direct as defence spending may involve opportunity costs by diverting resources away from healthcare and social sectors. This becomes especially relevant in developing countries where limited resources mean that higher defence expenditure may negatively impact human capital development. In contrast, in developed countries like the United States, the relationship tends to be positive, as higher defence spending often coexists with strong public investment in health, education and infrastructure.

These factors not only influence current health outcomes but also shape long-term human capital formation through improvements in innovation, institutional efficiency, and economic stability. Overall the study provides valuable insights into how defence spending interacts with human capital development highlighting the importance of balanced policy decisions to ensure sustained improvements in population health and overall economic well-being.

BIBLIOGRAPHY

- https://www.sipri.org/sites/default/files/2025-04/2504_fs_milex_2024.pdf
- https://towardfreedom.org/story/archives/americas/u-s-launched-251-military-interventions-since-1991-and-469-since-1798/
- https://news.un.org/en/story/2025/09/1165809
- http://article.sapub.org/10.5923.j.economics.20180805.03.html#Sec3
- https://journals.internationalrasd.org/index.php/joe/article/view/107/101
- https://www.sciencedirect.com/science/article/abs/pii/S0176268012000432
- https://fred.stlouisfed.org/series/SPDYNLE00INUSA
- https://fred.stlouisfed.org/series/FDEFX
- https://fred.stlouisfed.org/series/FYFRGDA188S
- https://fred.stlouisfed.org/series/SPPOPGROWUSA
- https://fred.stlouisfed.org/series/USAGFCFQDSNAQ
- https://journals.sagepub.com/doi/10.1177/0095327X261430680?int.sj-full-text.similar-articles.6
- https://www.imf.org/en/publications/wp/issues/2016/12/31/public-investment-public-finance-and-growth-the-impact-of-distortionary-taxation-recurrent-41518
- https://www.sciencedirect.com/science/article/abs/pii/S0167629613001306