ASSESSING THE EPIDEMIOLOGICAL TRANSITION THEORY'S UTILITY IN ENHANCING LIFE EXPECTANCY IN LEAST DEVELOPED COUNTRIES: A COMPARATIVE ANALYSIS

A THESIS

Presented to

The Faculty of the Department of Economics and Business

The Colorado College

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Arts

By:

Yousheng Tang

February 2024

ASSESSING THE EPIDEMIOLOGICAL TRANSITION THEORY'S UTILITY IN ENHANCING LIFE EXPECTANCY IN LEAST DEVELOPED COUNTRIES: A COMPARATIVE ANALYSIS

Yousheng Tang

February 2024

Economics & Mathematics

ABSTRACT

This study evaluates the applicability of the Epidemiological Transition Theory (ET) in addressing the contemporary challenge of improving life expectancy in Least Developed Countries (LDCs). Through the integration of classical and modern viewpoints within the discourse on changes in life expectancy, this study aims to uncover the causes behind global life expectancy disparities, with an emphasize on the impact of infectious diseases. Utilizing a multi-staged one-way fixed-effect ordinary least squares (FE-OLS) panel regression model, this analysis tests the hypothesis that infectious diseases disproportionately impact life expectancy in LDCs compared to other countries, particularly developed nations. The investigation employs comparative analysis based on different country classifications. The findings reveal significant determinants of life expectancy disparities, including the pronounced negative effect of HIV incidence and the ineffectiveness of Polio vaccine coverage in LDCs. The study highlights the ET's relevance and precision within the LDC context, despite the theory's criticisms. This analysis not only contributes to policy development aimed at enhancing well-being in LDCs but also reignites the ET as an instructive framework for understanding life expectancy challenges in the modern era.

KEYWORDS: LDC, Life Expectancy, Epidemiological Transition Theory

JEL CODES: I15, O15, J11

ON MY HONOR, I HAVE NEITHER GIVEN NOR RECEIVED
UNAUTHORIZED AID ON THIS THESIS

Signature

TABLE OF CONTENTS

Introduction

In 2015, all member states of the United Nations adopted the 2030 Agenda for Sustainable Development (United Nations, 2015). This agenda consists of 17 Sustainable Development Goals (SDGs) that aim to promote the prosperity of the world. Among these 17 blueprints for a more peaceful and sustainable future, SDG 3 in particular calls for ensuring healthy lives and promoting the well-being for all at all ages (the United Nations, 2015; Leboeuf, 2018). Life expectancy, as the focus of this study, is a broad measure of a nation's health status, which reflects the cumulative social economic, medical and technological achievements of the human society (Halicioglu, 2011; Liou, Joe, Kumar, & Subramanian, 2020).

Broadly speaking, the world at large witnessed important scientific breakthroughs and policy engagements that greatly advanced life expectancy in the last two centuries (Liou, et al., 2020). However, shifting focus from the general trend to the local context of underdeveloped regions reveals that the improvement in life expectancy is far from optimistic. Researchers found that many countries in the Eastern Mediterranean Region (EMR) noted lower life expectancies despite increases in income and health expenditures (Alwan & the World Health Organization (WHO), 2016). Even more concerning, life expectancies are still unacceptably low while mortality remains the opposite in Least Developed Countries (LDCs) today (the United Nations LDC 5, 2023). However, the comparative studies with a specific focus on the attribution of life expectancy disparity between LDCs and other countries and corresponding empirical models are extremely limited.

Since the beginning of the twentieth century, numerous scholars and researchers worldwide sought to elucidate factors that drive changes in life expectancy from various perspectives. These factors include but not limited to socio-economic factors

such as education and inequality (Novak, Čepar, & Trunk, 2016; Sen, 1991), environmental factors such as pollution (Preston, 1975; Sen, 1991), technological factors such as vaccination rates (Preston, 1975; Gilligan and Skrepnek, 2015). In 1971, Abdel Omran first introduced the famous Epidemiological Transition Theory (ET). The significance of the ET in the historical and modern evolution of mortality and life expectancy in the sociobiological context is unparalleled. No researcher in this field can avoid attempting to use the analytical and prognostic potential of the ET (Anatoly, 2020). In general, the ET describes the shifts in mortality and disease patterns and identifies three successive stages of such shifts (Anatoly, 2020; Omran, 1971). Specifically, the theory suggests that infectious diseases influence life expectancy differently in regions that are in different stages of epidemiologic transition (Omran, 1971). However, there is still a significant gap in exploring the practical implications of the ET in our time in the context of enhancing life expectancy in LDCs. One of the main reasons is that the ET failed to unify all theories and models within the study of the relationship between diseases and life expectancy and the criticism targeting its completeness and practicability continuous today.

In 1975, Preston systematically analysed the substantial increase in life expectancy in the 20th century with an emphasis on the contribution of national income level to life expectancy. Preston demonstrated that the relationship between life expectancy and national income level follows a logarithmic curve and this curve was also known as the Preston Curve (Preston, 1975). Although Preston's finding aligned with Omran's view that economic advance is not the primary and direct factor of life expectancy, Preston distinctively expressed a different opinion than Omran regarding the different influences of infectious diseases on life expectancy (Preston, 1975; Omran, 1971; Omran 1998). Preston pointed out that infectious diseases and

corresponding vaccine coverage, as factors exogenous to a country's income level, had similar effects on life expectancies across different level of economic development, in both developed and less developed countries (Preston, 1975). In 2014, another study found that the ET's theoretical approaches are unsuitable for analysing the demographic situation in Africa (Defo, 2014). Additionally, many developing countries are suffering under a "double burden of disease" that was not fully explained by the ET (Nnebue, 2010).

This analysis investigates the feasibility of utilizing the Epidemiological Transition Theory in answering the contemporary challenge of enhancing life expectancy in LDCs. This study combines both classical and modern perspectives within the discourse on life expectancy changes, aiming to establish the causes of global life expectancy disparity and illuminate the reasons for the low life expectancy in LDCs with an emphasis on the influence of infectious diseases. To do so, a one-way fixed-effect ordinary least squares (FE-OLS) panel regression model is constructed to examine the hypothesis: infectious diseases have a disproportionately larger negative impact on life expectancy in Least Developed Countries compared to other countries, especially developed nations. Existing studies often overlook the distinct characteristics of LDCs by broadly categorizing them under developing countries, without conducting comparative analyses based on varied country classifications in FE estimation methodologies. Similarly, only a limited number of modern theories utilize large time-span panel datasets, with LDCs as the focus, to investigate the accuracy and feasibility of the ET. This study aims to bridge the gap in existing literature by utilizing the ET as a guiding principle to identify the reasons for low life expectancy through comparison with other categories of countries. Through the examination of the hypothesis, this analysis aims to contribute valuable insights to

policy development in LDCs, and guide efforts to address health disparities around the world.

Literature Review

The following section is divided into three parts. The first part shows the global trend of life expectancy. The second part meticulously elucidates the Epidemiological Transition theory and its relevance to this analysis. The third part explains the different impact of vaccination programs on life expectancy across different regions.

Global Life Expectancy Overview

In 1800, no region around the world had a life expectancy higher than 40 years (Roser, Ortiz-Ospina, & Ritchie, 2013). However, by 2021, the average life expectancy surpassed the 70-year mark (Roser et al., 2013). Additionally, by examining data from eight nations presently recognized as developed nations, Oeppen and Vaupel revealed that both male and female life expectancy rose in a stunningly linear pattern from 1840 to 2000 (Oeppen & Vaupel, 2002). This trend is corroborated by another study, extending the linear pattern of life expectancy to 2012 (Burger, Baudisch, & Vaupel, 2012). In the longest-lived national populations, life expectancy increased by about 3 months per year (Burger et al., 2012). Although the disparity in life expectancy at birth steadily reduced over time between the best performing and the worst performing regions, the gap remains disquieting (Liou, et al., 2020). From 2010-2015, the gap was 35.3 years between Lesotho, an African country still recognized as a LDC by the United Nations in 2023, and China, Hong Kong SAR (the United Nations LDC5, 2023; Liou, et al., 2020). This concerning revelation implies that a child born in Lesotho between 2010 and 2015 was likely to experience a life expectancy of almost a half (49.5%) the global average of 71.3 years (Lious, et al., 2020).

The Epidemiological Transition Theory

Core Concept. The ET, as the foundational principle that inspired this analysis, deserves the primary attention for a thorough examination and interpretation of its core viewpoints, limitations, and appropriate applications. It is noteworthy that Omran actually published two articles about the ET. The first one had wider recognition, extensive discussion, expansion, and criticism. Omran drafted the second article shortly before his death and his colleagues completed the article (Omran, 1998). The second article primarily functioned as Omran's response to three decades of critique and extension on the ET, introducing two additional stages to the original theory published in 1971. For the purpose of investigating the enhancement of life expectancy in the LDCs, the more informative source is Omran's first article because LDCs remain distant from the additional two final stages outlined in the second article. Despite how later scholars may extend or refine the stages included in the ET, its core viewpoint remains constant: the assertion of the integral and natural process that human society transitions from the era of infectious diseases to the era of chronic diseases accompanied by a reduction in mortality (Omran, 1971; Anatoly, 2020). According to Omran in 1971, this inevitable transition process can be categorized into three stages: the age of pestilence and famine, the age of receding pandemics, and the age of degenerative and man-made diseases. In fact, from a worldview, Omran's view of the whole panorama of diseases patterns changes is broader and more significant than most of the views of his critics (Anatoly, 2020). For example, the supplementary fourth stage, the "delayed degenerative diseases stage", proposed by Olshansky and Ault in 1986 can still be fully covered by the third stages within the ET 1971 because delayed degenerative diseases remain degenerative (Olshansky and Ault, 1986).

Limitations and Scope of Application. The ET demonstrated its inaccuracy in predicting the future trends of infectious diseases and their impact on life expectancy in developing countries in Africa. In 2014, Defo found that infectious and parasitic diseases alone account for 41% of all deaths in Africa, more than the double of the prediction by the ET, which suggested a decline to 19.4% in 2015 (Defo, 2014; Omran, 1998). While currently there is not a similar study in the context of LDCs, it can reasonably be assumed that the ET's predictions for the future life expectancy in LDCs will also be highly inaccurate. There are two main reasons: firstly, Defo's findings in Africa are highly pertinent since most LDCs today are African countries and very few countries can break free from the category of LDC; secondly, Omran's first article on the ET was published around the same time as the UN's official establishment of the LDC category, and Omran's construction of the ET dates back to the mid-1960s, implying that he did not specifically consider the special category of the countries later known as the LDCs (Defo, 2014; Fialho, 2012; Omran, 1971; Anatoly, 2020; Omran; 1998).

Although these facts highlight the significant limitations of the ET in predicting life expectancy in LDCs, it does not diminish the application and relevance of the ET in addressing the hypothesis of this study. The primary focus of this research is to attribute the low life expectancy in LDCs to specific and systematic causes rather than predict its future trends. The ET and Omran's explanation on changes in life expectancy demonstrated incredible insights and accuracy far beyond its time. In Omran's 1971 article, he explicitly state: "the reduction of mortality in most western countries during Europe in the nineteenth century, as described by the classical model of epidemiologic transition, was determined primarily by ecobiologic and socioeconomic factors. The influence of medical factors was largely inadvertent until

the twentieth century" (Omran, 1971). People only started accepting this viewpoint in recent years. "Nowadays, the prevailing point of view is that immunization and effective treatment methods have borne fruit only in the 20th century" (Livi Bacci, 2010, as cited in Anatoly, 2020). The practicality and accuracy of the ET in the attribution of causes of changes in life expectancy made it possible and necessary to the verification of the hypothesis. The subsequent step is evaluating the stages to which all countries, including LDCs, belong. This assessment aims to provide the most fitting evaluation of the applicability of the ET in LDCs.

Stages Clarification. Firstly, the LDCs are in an overlapping state between the first and second stages of the ET, a condition previously mentioned as "double burden of diseases" (Nnebue, 2010). The simultaneous existence of multi-stages of the transition indicates the incompleteness of the transition into the age of degenerative diseases in LDCs (Anatoly, 2020). In Omran's words from the second article of the ET, this condition is caused by the combination of "unfinished old health problems" and "rising new health problems" (Omran, 1998). Numerous studies indicate that Africa is experiencing an overlapping state, particularly with the prevalence of infectious diseases such as HIV/AIDS, diarrhea, and malaria, along with noncommunicable diseases such as cardiovascular diseases, cancer, and obesity (Boutayeb, 2010; Frenk & Gomez-Dantes, 2011). By 2019, out of the world's 48 least developed countries, 34 are in Africa, which directly transmitted the overlapping status to the LDCs as a whole (Anatoly, 2020). Developed countries completed the epidemiological transition, meaning the significant reduction of burden of infectious diseases and the ongoing of the age of degenerative and man-made diseases. These characteristics are most evident in G7 countries as they represent the most developed economies in the world and completed the transition early. According to Omran,

western developed countries led by the UK gradually began the transition as early as 1920 while cardiovascular deaths largely increased in 1945 (Omran, 1971). However, clarifying a stage for developing countries (excluding LDCs) uniformly is difficult because of the uneven progress in their epidemiological transition. For example, although Nigeria is not classified as an LDC, it experiences a condition similar to LDCs with a severe overlapping state: there is an "increasing prevalence of lifestyle-associated diseases like diabetes mellitus, and of overweight and obese individuals, as well as the dominance of infectious diseases" (Nnebue, 2010). Meanwhile, similar to Omran's prediction in the second article of the ET, China, also a developing country, completed the transition (Omran, 1998). In 2018, over 86% of deaths were caused by Non-communicable Diseases (NCD) in China (the WHO, n.d.-a). Therefore, the refinement of the investigation of the hypothesis is equivalent to the demonstration of a gradual decrease in infectious diseases' average impact on life expectancy in LDCs, developing countries (excluding LDCs), developed countries, and G7 nations.

HIV/AIDS. HIV/AIDS is the abbreviate for Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome. HIV is a virus that attacks the body's immune system and it can lead to AIDS if HIV is not treated (the Centers for Disease Control and Prevention, 2022). Since its first discovery in 1983, HIV infection is a major global health issue, affecting 36.7 million people world-wide (Barré-Sinoussi, Chermann, Rey, Nugeyre, Chamaret, Gruest, ..., & Montagnier, 1983; Hsu & O'Connell, 2017). Even more disquieting, HIV infections is still incurable and none of the numerous attempts over many years to develop an HIV vaccine convincingly succeeded (D.Y.Lu, T.R.Lu, Zhu, Yarla, Che, & Wu, 2017; Burton, Ahmed, Barouch, Butera, Crotty, Godzik, ..., & Wyatt, 2012). Antiretroviral therapy (ART) is the major method to control HIV infections and the number of people living with HIV on ART

reached 17 million in 2015 (Hsu & O'Connell, 2017). ART can dramatically reduce morbidity and mortality in individuals with HIV infection but the need for ART is lifelong and the cost is substantial and difficult to sustain economically (Hsu & O'Connell, 2017; Deeks, Lewin, Ross, Ananworanich, Benkirane, Cannon, ... & Zack 2016).

Vaccination

Effectiveness. It is a common view that vaccination is one of the greatest health achievements and most effective preventive health measures against infectious diseases to date (Gilligan & Skrepnek, 2015). The progress and widespread adoption of medical technology have an undeniable positive impact people's well-being around the world. The focus of this analysis is whether vaccine technology, due to its application in different countries, has varying effects on health, measured by life expectancy. Similar to the discussion on infectious diseases, the ET could not and did not specifically mention the implementation of immunization in developing countries due to its temporal constraint. The first large implementation of vaccine in developing countries occurred during the Expanded Programme on Immunization (EPI) launched by the WHO in 1974, which aimed to ensure that all children in all countries benefited from life-saving vaccines (the WHO, n.b.-b). However, a correct interpretation of the ET can still imply Omran's answer to this question. In 2020, Anatoly analyzed that Omran's concept of the epidemiological transition actually includes the "sanitary transition", "health transition", and "the patterns of the organized social response to health condition" (Omran, 1971; Anatoly, 2020). As part of sanitation and the social response to infectious diseases, vaccination is expected to have effects that change with the evolution of stages on life expectancy.

Preston questioned Omran's viewpoint on this problem in 1975 and presented his argument. Preston claimed that vaccination, as a health technology and a factor exogenous to income level, should equally influence health conditions and life expectancies in both developed and developing countries because "the techniques, once proved effective, spread rapidly from country to country" (Preston 1975). However, research showed that Preston's inference on vaccination is incorrect, because the application and spread of vaccines in developing countries cannot directly infer their effectiveness. By analyzing data from the Eastern Mediterranean Region between 1995 and 2010, Gillian and Skrepnek found that the growth in vaccination coverage in Egypt, Iran, Iraq, Libya, Morocco, Oman, Saudi Arabia, Syria, and Tunisia had a significantly higher average positive effect on life expectancy compared to its effects in Afghanistan and Somalia, which are still classified as LDCs today (Gilian & Skrepnek, 2015; the United Nations LDC 5, 2023). This fact once again confirms the foresight and validity of the ET.

In conclusion, despite significant increases in life expectancy worldwide, disparities persist, with LDCs experiencing lower life expectancies compared to other countries. The ET, while facing limitations in predicting life expectancy in LDCs, remains relevant in identifying systematic causes of low life expectancy. Omran's theory provides valuable insights into the transition from infectious to chronic diseases, guiding the investigation of low life expectancy causes. Additionally, the effectiveness of infectious disease vaccinations varies across regions, which implies socio-economic disparities and aligning with the principles of the ET. The literature underscores the pressing need for understanding life expectancy disparities fully through a reliable approach. The following section explicitly explains the empirical models of this analysis.

Empirical Model

This analysis employs an extended version of the semi-logarithmic model used by Halicioglu for modelling life expectancy in Turkey (Halicioglu, 2011). This study modified the model to accommodate the panel data framework. Halicioglu categorized the determinants of life expectancy into three groups: economic factors, social factors, and environmental factors. In addition to Halicioglu, numerous theories supported and further extended these three groups of factors that affect life expectancy.

Income Level. In 1975, Preston systematically analysed data from the 1900s, 1930s, and 1960s and identified a logarithmic curve as the relationship between life expectancy and income level (Preston, 1975). Other researchers call this curve the Preston Curve. Preston also emphasized that the impact of income level on life expectancy is indirect, only directly influencing the consumption of health-related items such as food, housing, leisure, and products affecting health negatively, like cigarettes (Preston, 1975). While the Preston Curve originally described the relationship between national income per capita and life expectancy, some studies use GDP per capita as a substitute for national income per capita when referring to the Preston Curve. For instance, Felice, Andreu, and D'Ippoliti applied the Preston Curve in their research to life expectancy in Italy and Spain in 2016 and they explained that "the relationship between GDP and life expectancy in the 20th century follows a logarithmic curve" and "the impact of GDP on life expectancy is higher when the former is low, then it decreases as GDP rises, and even disappears after GDP reaches a certain threshold" (Felice, Andreu, & D'Ippoliti, 2016). This analysis builds upon the relationship between income level and life expectancy uncovered by Preston, incorporating the GDP per capita of a country as an independent variable.

Thinness and Obesity. Thinness caused by undernutrition is prevalent in the overall population of developing countries, particularly among adolescents in some countries, and is one of the major health problems faced by developing countries (Mushtaq, Gull, Khurshid, Shahid, Shad, & Siddiqui, 2011). Back in 2014, researches used multi-staged random sampling in Kano, Nigeria and found that the overall prevalence of thinness was 60.6% among the selected students from six secondary schools (Mijinyawa, Yusuf, Gezawa, Musa, & Uloko, 2014). Body Mass Index (BMI) is one of the most commonly used standards for thinness and obesity. BMI is a value derived from the mass (weight) and height of a person. The BMI is defined as the body mass divided by the square of the body height, and is expressed in units of kg/m^2, resulting from mass in kilograms (kg) and height in metres (m) (Wikipedia, 2023). Numerous studies show that BMI is a significant factor of people's well-being. The classifications of adult BMI are underweight (<18.5 kg/m^2), normal weight (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), obese (30-40 kg/ m^2), and severe obese (over 40 kg/m^2) (Wikipedia, 2023; Walls, Backholer, Proietto, & McNeil, 2012). Obesity and overweight are associated with large decreases in life expectancy (Peeters, Barendregt, Willekens, Mackenbach, Mamun, Bonneux, & NEDCOM, 2003). Another study in 2016 that used meta-analysis of 239 prospective studies in four continents showed mortality was minimal at the BMI range 20.0-25.0 kg/m^2 and increased significantly below this range and throughout the overweight range (Di Angelantonio, Bhupathiraju, Wormser, Gao, Kaptoge, De Gonzalez, ..., & Hu, 2016). This analysis incorporates BMI and thinness as independent variables that affect life expectancy.

Alcohol Consumption. Alcohol consumption is a major correlate of health and disease. Numerous studies showed that alcohol drinking increases the risk of cancers

including oral cavity, pharynx, and larynx (Corrao, Bagnardi, Zambon, & La Vecchia, 2004). This analysis includes the alcohol consumption of a country as an independent variable of life expectancy due to its association with man-made and life-style diseases.

Inequality. In 1991, Sen explained how inequality significantly reduce life expectancy even when national income level is high. Combining the severe racial inequalities in the United States at the time, Sen proposed that inequalities in health services, social security, and environmental conditions caused by "inadequacy in market mechanisms in efficiently generating and equitably distributing some of the basic requirements of good living" led to remarkably low life expectancy among African Americans (Sen, 1991). Thus, inequality led to lower average life expectancy in the United States. This analysis incorporates inequality as one of the independent variables affecting life expectancy, using the Gini Index, a measure of wealth inequality within a nation or a social group, as the standard for assessing inequality levels (the World Bank, 2024).

Political Philosophy. Sen also claimed that, in terms of enhancing life expectancy, “the capitalist economies do much better than the socialist ones” (Sen, 1991). Sen further explained that the political philosophy of a country affects the life expectancy through environmental factors. State-owned industries and public firms of socialist economies are more destructive to the environment because of their obtuseness toward public criticism and environmental needs (Sen, 1991). This study incorporates the political philosophy of a country as an independent variable into consideration.

Education. In 2016, through multiple regression analysis on cross-section data of 187 countries around the world, Novak, Čepar, and Trunk found that "a country's education level is an important determinant of life expectancy at birth" (Novak et al., 2016). Based on data of year 2010, the researchers found that the expected years of schooling has a positive effect on life expectancy at birth (Novak et al., 2016). This study incorporates the years of school of a country as an independent variable into consideration.

Vaccines. This analysis also focuses on vaccines for Hepatitis B, Measles, Poliomyelitis (Polio), and Diphtheria and investigate their effect on life expectancy. Around 33% of the world population has past or present infection by the Hepatitis B Virus (HBV) which shows a significant prevalence in South-East Asia and Sub-Saharan Africa (Marcellin, 2009). HBV can lead to chronic hepatitis, cirrhosis, and hepatocellular carcinoma, with maternal-to-infantile transmission being the major transmission (Te & Jensen, 2010). Infantile vaccination is the primary method to control Hepatitis B (Alavian, 2011). Measles, Polio, and Diphtheria, like HBV, are all significant challenges to public health, especially in underdeveloped regions, and the corresponding vaccines serve as one of the most effective measures against these infectious diseases (Assaad, 1983; M. M. Mehndiratta, P. Mehndiratta, & Pande, 2014; Sharma, Efstratiou, Mokrousov, Mutreja, Das, & Ramamurthy, 2019).

Following the literature, the average life expectancy of both genders can be written as a function f:

$$L = f(E, I, H, S, C, EN) \tag{1}$$

L: The average life expectancy.

E: Economic Factors, represented by GDP per capita (GDP) in this analysis.

I: Infectious Diseases Related Factors, represented by HIV/AIDS incident (HIVI), Hepatitis B (HB), Measles (M), Polio (P), Diphtheria (D) vaccine coverage rates.

H: Health Condition Related Factors, represented by the prevalence of thinness of children between age 5-9 (TC), the prevalence of thinness of adolescents between age 10-19 (TA), and body mass index (BMI).

S: Social Factors, represented by schooling years (SY) and Gini index (GI).

C: Chronic Diseases Related Factors, represented by alcohol consumption (AC).

EN: Environmental Related Factors, represented by political philosophy (PP)[1]

In order to minimize time-related heterogeneity and the heterogeneity of different epidemiological transition stages, this analysis divides the research subjects into three groups, LDCs (group 1), developing countries excluding LDCs (group 2), and developed countries (group 3). When investigating the hypothesis within group 1 and group 2, this analysis applies the following empirical model generated from function (1):

[1] Instead of using crime rate per capita and urbanization to reflect living environment, this analysis adopts Sen's conclusion, using PP as a factor to reflect natural environment that more closely related with pollution and sustainable development.

$$L_{it} = GDP_{it}\beta_1 + HIVI_{it}\beta_2 + HB_{it}\beta_3 + M_{it}\beta_4 + P_{it}\beta_5 + D_{it}\beta_6 + TC_{it}\beta_7 + TA_{it}\beta_8 + BMI_{it}\beta_8 + SY_{it}\beta_9 + GI_{it}\beta_{10} + AC_{it}\beta_{11} + PP_{it}\beta_{12} + \alpha_i + \epsilon_{it} \quad (2)$$

where L is the average life expectancy of both genders of a country i in year t; β_{1-12} are coefficients of the corresponding independent variables. The variable α_i represents a country's unobserved FE which includes culture, geographic location, etc. The variable α_i captures country-specific heterogeneity that is irrelevant to time. The variable ϵ_{it} is an idiosyncratic error term and its expected value given all independent variables is 0. Notice that α_i can effectively capture unobserved heterogeneity in group 1 and group 3 because there is no stage-specific heterogeneity within these two groups. As mentioned earlier, the LDCs find themselves in the overlapping phase of the first two epidemiological stages, while developed countries completed the transition. A country's economy, disease patterns, environmental situation, health condition, and social climate are closely related with individual unobserved heterogeneity. In this case, $cov(E, \alpha) \neq 0, cov(I, \alpha) \neq 0, cov(H, \alpha) \neq 0, cov(S, \alpha) \neq 0, cov(C, \alpha) \neq 0, cov(EN, \alpha) \neq 0$. This analysis assumes that both LDCs and developed countries remain within their respective epidemiological stages without transitioning to other stages during the observed period. Additionally, it assumes that within these stages, the countries do not exhibit time-varying heterogeneity in their epidemiological characteristics. The two assumptions enable the one-way FE-OLS panel regression model to capture all unobserved heterogeneity in group 1 and group 3 countries.

However, the same methodology would not work for group 2 since it is a fact that some developing countries finished the epidemiological transitions, such that

model (2) is unable to capture the stage-specific heterogeneity that leads to biased estimation. To solve this problem, an extended model based on function (2) is constructed for group 2:

$$L_{itj} = GDP_{it}\beta_1 + HIVI_{it}\beta_2 + HB_{it}\beta_3 + M_{it}\beta_4 + P_{it}\beta_5 + D_{it}\beta_6 + TC_{it}\beta_7 + TA_i\beta_8 + BMI_i\beta_8 + SY_{it}\beta_9 + GI_{it}\beta_{10} + AC_{it}\beta_{11} + PP_{it}\beta_{12} + \alpha_i + \delta_j + \epsilon_{ijt} \quad (3)$$

where a new variable δ_j is added to capture the stage-specific heterogeneity; L now represents the average life expectancy of both genders of a country i in year t with epidemiological transition stage j.

As explained by the ET, the patterns of infectious diseases and chronic diseases within a country will change once the transition happens. In this case, $cov(I, \delta) \neq 0 \; and \; cov(C, \delta) \neq 0$, which cannot be removed by one-way FE estimation. As inspired by the empirical strategy used by Xiao, et al., this analysis applies the proxy variable (PV) approach to mitigate the endogeneity caused by the unobserved stage-related characteristics (Xiao, Li, & Fleisher, 2015). Identifying a specific epidemiological stage for every country requires multifaceted and intricate examination and is beyond the scope of this analysis. This analysis uses three proxy variables to mitigate the correlation between unobserved stage-related heterogeneity and disease patterns. The model for group 2 is an extended version of function (3):

$$L_{it} = GDP_{it}\beta_1 + HIVI_{it}\beta_2 + HB_{it}\beta_3 + M_{it}\beta_4 + P_{it}\beta_5 + D_{it}\beta_6 + TC_{it}\beta_7 + TA_{it}\beta_8 + BMI_{it}\beta_8 + SY_{it}\beta_9 + GI_{it}\beta_{10} + AC_{it}\beta_{11} + PP_{it}\beta_{12} + \alpha_i + AG_{it} + ID_{it} + UFD_{it} + \epsilon_{it} \quad (4)$$

where AG, ID, UFD are three dummy variables; AG takes the value of 1 if the average life expectancy of a developing country i in year t exceeds 70 years, a key indicator of the completeness of epidemiological transition, and 0 otherwise; ID returns 1 if the number of infant deaths per 1000 population of a developing country i in year t is less than or equal to the average value of the infant deaths per 1000 population of all developed countries in the same year t, and 0 otherwise; UFD returns 1 if the number of deaths of children under five years old per 1000 population of a developing country i in year t is less than or equal to the average number of deaths of children under five years old per 1000 population of all developed countries in the same year. Similar to the characteristics of the third stage describe by the ET, these three proxy variables provide valuable indications of whether a developing country is in the degenerative and man-made diseases age.

To summarize, this analysis employs model (2) for investigating the relationship between the average life expectancy and its determinants within LDCs and all developed countries and model (4) for developing countries. Additionally, several modifications are necessary for model (2) and model (4) to get robust result after investigating the correlations among the variables. The details of the modifications and the final models are in the following section after the explanation of the data used in this analysis, the descriptive statistics, and the analysis of the distribution of the dependent variable.

Data

The data used in this analysis is from multiple sources. The Global Health Observatory (GHO) data repository under the WHO collected the health status and heath related factors used in this analysis. In 2023, a data scientist from the Kaggle platform collected and modified the original data set of the WHO (Kaggle, 2023). After modifications, the dataset used by this analysis contains life expectancy, health, economic and demographic information for 179 countries from year 2000-2015, for a total of 2864 observations.

Variables

This analysis includes 21 variables and uses average life expectancy as the dependent variable because it captures important components of health status. Among these 21 variables, 16 variables are from the original dataset:

Average Life Expectancy (L): Average life expectancy of both genders.

Country: The names of 179 countries used in this analysis. There are 44 LDCs, 97 developing countries (LDCs excluded), and 38 developed countries[2].

Year: 16 years observed from 2000-2015.

GDP per capita (GDP): GDP in current USD.

HIV incident (HIVI): Incidents of HIV per 1000 population aged 15-49.

[2] The recognition of LDCs is based on the latest information from the UN. Sudan, South Sudan, and Tuvalu are excluded due to missing data. Maldives graduated from the LDC category in 2011 but this analysis treats it as an LDC through the observed period in order to keep the panel dataset balanced.

Hepatitis B vaccine coverage rate (HB): Hepatitis B (HepB3) immunization coverage rate (%) for 1-year-olds.

Measles vaccine coverage rate (M): Measles immunization (containing first does) coverage rate (%) for 1-year-olds.

Polio vaccine coverage rate (P): Polio immunization coverage rate (%) for 1-year-olds.

Diphtheria vaccine coverage rate (D): Diphtheria immunization coverage rate (%) for 1-year-olds.

Thinness among five-nine years old children (TC): Prevalence (%) of thinness among children aged 5-9 years (BMI < -2 standard deviations below the median).

Thinness among ten-nineteen years old adolescents (TA): Prevalence (%) of thinness among adolescents aged 10-19 years (BMI < -2 standard deviations below the median).

Body mass index (BMI): The average BMI of individuals within a country. BMI is a measure of nutritional status in adults. It is defined as the body mass divided by the square of the body height.

Schooling year (SY): Average years that people aged 25+ spent in formal education.

Alcohol consumption (AC): Liters of pure alcohol consumption per capita aged 15+ years old.

Infant death (InfantD): Infant deaths per 1000 population.

Under-five death (UnderFD): Deaths of children under five years old per 1000 population.

Other 5 variables are generated by author and merged into the original dataset. As outlined above, the analysis of group 2 countries requires three proxy variables to mitigate the stage-related heterogeneity. These three dummy variables are:

Average age over 70_dummy variable (AG): This variable determines whether the average life expectancy of a specific observation is above 70 years old (returns 1 if it is true).

Infant death_dummy variable (ID): This variable determines whether the value of InfantD of a specific observation is lower than the average of the InfantD of 38 developed countries in the same year (returns 1 if it is true).

Under-five death_dummy variable (UFD): This variable determines whether the value of UnderFD of a specific observation is lower than the average of the UnderFD of 38 developed countries in the same year (returns 1 if it is true).

Additionally, this analysis adds another 2 variables according to the literature to access the inequality level and political philosophy of a specific country:

Political philosophy_dummy variable (PP): The political philosophy of a country (returns 1 if the country is a socialist state)[3].

[3] Among the 179 countries included in this analysis, 15 of them are Marxist-Leninist states or Non-Marxist-Leninist states with constitutional references to socialism (Wikipedia, 2024).

Gini Index (GI): The Gini index, also known as the Gini coefficient, of a country[4].

Many countries, especially LDCs, lack consistent Gini index statistics, and some have no available data within the observed timeframe from 2000 to 2015. To improve data accessibility and maintain a balanced panel dataset, this study collected the Gini index of year 2019 for each country as a proxy for the Gini index from 2000 to 2015, assuming GI is time-invariant. In cases where a country lacked an available 2019 Gini index, this analysis utilizes the nearest available Gini index to 2019. Table 4.1 in Appendix A shows the multiple sources of the Gini index for each country.

Descriptive Statistics. Table 4.2.1 presents the means and dispersions of 17 numerical variables used in this analysis across three groups of countries. Through horizontal comparison, Table 4.2.1 shows that, between 2000 and 2015, the average life expectancy of LDCs is approximately ten years lower than that of developing countries, which is similar to the gap between the latter and the developed countries. The average life expectancies of the three groups of countries aligns with the inference of the ET, which states that 70-years of average life expectancy is a significant indicator of the completion of epidemiological transition (Omran, 1971). Within the observed timeframe, the average life expectancy of LDCs is only 58.5 years, while that of developed countries is close to 80 years. The average life expectancy of developing countries is very close to 70 years, which is consistent with the fact that some developing countries completed the epidemiological transition.

The three groups also exhibit significant differences in terms of HIV and vaccine coverage rates. On average, the prevalence of HIV in LDCs is approximately 18 times

[4] The Gini index is a measure of the income inequality within a nation.

higher than in developed countries and 0.5 times higher than in developing countries within the observed timeframe. In regard to the vaccine coverage rates of four infectious diseases considered by this analysis (Hepatitis B, Polio, Measles, and Diphtheria), LDCs notably fall behind both group 2 and group 3 countries. Measles vaccine coverage rate is notably lower in LDCs, standing at 64%, compared to 88% in developed countries. Furthermore, the standard deviations of coverage rates for Polio and Diphtheria vaccines in LDCs surpass the standard deviations seen in both group 2 and group 3 countries. This suggests an imbalance in the allocation of medical resources within LDCs.

TABLE 4.2.1

DATA SUMMARY

	LDCs		Developed		Developing	
Variable	Mean	SD	Mean	SD	Mean	SD
L	58.542	7.021	78.559	3.144	69.733	7.151
HB	75.842	17.188	88.061	14.394	86.649	14.621
M	64.621	14.923	88.434	9.977	78.773	19.47
BMI	22.571	1.615	26.011	1.003	25.767	1.908
P	74.314	18.436	94.9	3.977	88.736	12.628
D	74.337	18.77	95.107	3.78	88.224	13.457
HIVI	1.45	2.555	.076	.044	.963	2.665
GDP	1078.972	1259.699	33972.313	21143.756	7498.999	9509.436
TA	8.277	4.48	1.272	.734	4.727	4.085
TC	8.233	4.569	1.245	.802	4.82	4.219
SY	3.62	1.423	11.332	1.302	8.003	2.043
AC	2.135	2.425	9.832	2.799	4.076	3.141
GI	48.176	9.207	66.268	8.42	51.833	10.682
PP	.136	.343	.026	.16	.082	.275
AG	N/A	N/A	N/A	N/A	.628	.484
ID	N/A	N/A	N/A	N/A	.004	.062
UFD	N/A	N/A	N/A	N/A	.002	.044
Number of Observations:	704		608		1552	

Source: Generated by author.

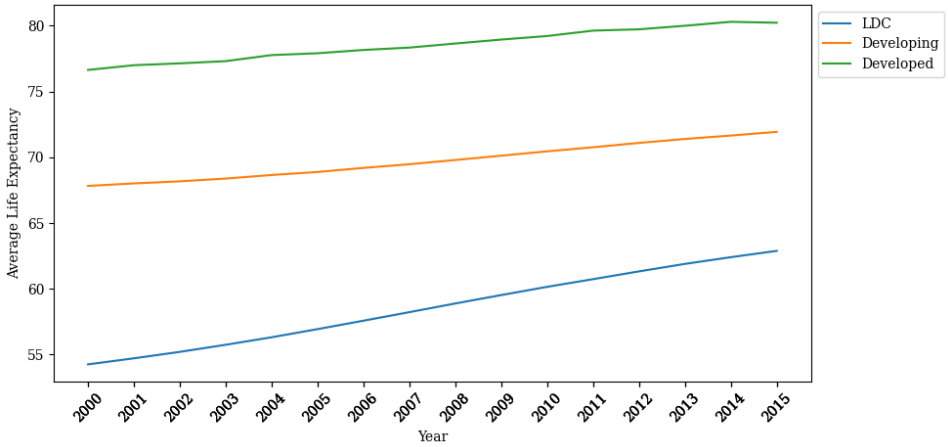
Data Distribution

This subsection meticulously analyzes the distribution of the dependent variable, average life expectancy, across three distinct groups of countries and throughout the observed timeframe from 2000 to 2015. Figure 4.3.1 illustrates the

trends in average life expectancies across the groups from 2000 to 2015. It indicates a consistent increase in average life expectancies within each group during this period, with LDCs experiencing the most rapid growth. The narrowing gap between these groups suggests a convergence in global life expectancy. This observation is consistent with the findings of Lious et al. in 2020, who documented a steady decline in the disparity of life expectancy worldwide in the last two centuries.

FIGURE 4.3.1

AVERAGE LIFE EXPETANCY



Source: Generated by author.

Additionally, Figure 4.3.2, Figure 4.3.3, and Figure 4.3.4 in Appendices A and B offers valuable insights into the evolving patterns of average life expectancy distribution among three groups of countries separately from 2000 to 2015. As the figures show, there is a discernible trend towards negative skewness in the distribution of life expectancy across these groups. This trend is consistent with Kannisto's observations regarding the shifting trends in life expectancy within developed nations throughout the 20th century, where the mode increases while the right-hand slope approaches a vertical line, as it was meeting “an invisible wall”

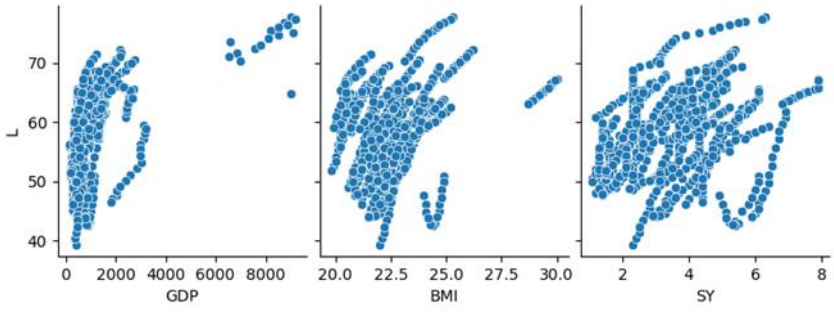
(Kannisto, 2001). Moreover, the trend is especially pronounced within group 1, the LDCs, which initially exhibited a positively skewed distribution of average life expectancy, with the mode lagging behind the mean in 2000. Within just 16 years, this distribution swiftly transitioned into a negatively skewed pattern. While the changes in LDCs are the most noticeable, as of 2015, the negative skewness in the LDCs group still remains lower than that of the other two groups[5].

Modifications

Several modifications are necessary for group 1 model. First of all, this analysis incorporates insights from the Preston Curve, applying a logarithmic transformation for both the dependent variable L and the independent variable GDP to linearize their relationship. Two additional independent variables, BMI and SY, also underwent logarithmic transformation due to their demonstrated logarithmic relationship with L shown in Figure 4.4.1.

FIGURE 4.4.1

LDCs PAIRPLOT (L, GDP, BMI, SY)



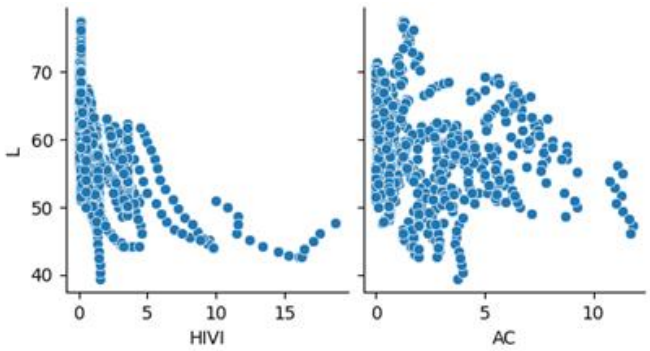
Source: Generated by author

[5] This analysis applies Pearson's First Skewness Coefficient, which is defined as (Mean-Mode)/Standard Deviation (Kovchegov, 2022). The coefficients of the three groups are: -0.34 for group 1; -0.58 for group 2; -0.40 for group 3.

Figure 4.4.2 reveals an inverse relationship between the independent variables $HIVI$ and AC with L. $HIVI$ and AC underwent an inverse transformation.

FIGURE 4.4.2

LDCs PAIRPLOT (L, HIVI, AC)

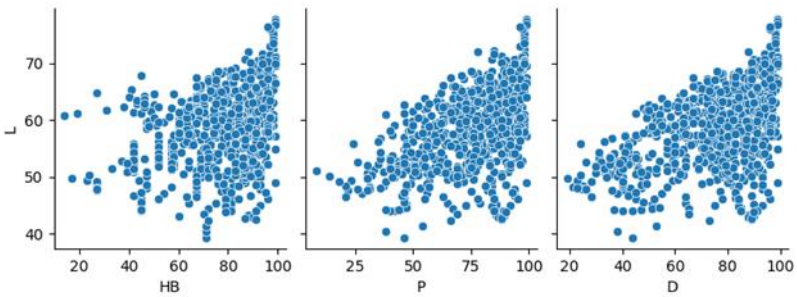


Source: Generate by author

Similarly, Figure 4.4.3 demonstrates a nonlinear relationship between HB, P, D and L. This analysis applies a square root transformation to HB, P and D.

FIGURE 4.4.3

LDCs PAIRPLOT (L, HB, P, D)



Source: Generated by author

Lastly, in order to counteract the logarithmic transformation of the dependent variable L, $HIVI$, AC, HB, P and D underwent a logarithmic transformation. A modification is also made to the observations of AC that has 0 values. An artificial value 0.0001 is then applied to all 0 values in order to make the inverse transformation possible and also make the dataset strongly balanced.

Table 4.4.1 presents the correlation matrix of the 13 independent variables for group 1 countries after the transformations.

TABLE 4.4.1

CORRELATION MATIRX (GROUP 1)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) lnGDP	1.000												
(2) lnBMI	0.494	1.000											
(3) lnSY	0.348	0.407	1.000										
(4) lnreAC	-0.075	-0.003	-0.218	1.000									
(5) lnreHIVI	0.255	-0.020	-0.071	0.584	1.000								
(6) lnsqrtHB	0.231	0.172	0.311	-0.277	-0.041	1.000							
(7) lnsqrtP	0.232	0.208	0.290	-0.194	0.058	0.655	1.000						
(8) lnsqrtD	0.341	0.243	0.379	-0.237	0.052	0.756	0.910	1.000					
(9) GI	-0.357	-0.264	-0.384	0.236	0.037	-0.110	-0.036	-0.114	1.000				
(10) PP	0.013	-0.224	0.127	0.028	0.119	0.099	0.171	0.167	0.170	1.000			
(11) TA	0.061	-0.295	-0.189	0.265	0.316	0.013	0.091	0.065	0.275	0.156	1.000		
(12) TC	0.096	-0.282	-0.151	0.256	0.326	0.027	0.103	0.092	0.253	0.170	0.864	1.000	
(13) M	0.322	0.055	0.085	-0.053	-0.003	0.005	0.129	0.127	-0.113	0.110	0.053	0.078	1.000

Note: lnGDP = ln(GDP); lnBMI = ln(BMI); lnSY = ln SY; lnreAC = ln($\frac{1}{AC}$); lnreHIVI = ln($\frac{1}{HIVI}$); lnsqrtHB = ln $\sqrt{HB}$; lnsqrtP = ln $\sqrt{P}$; lnsqrtD = ln $\sqrt{D}$.

Source: Generated by author

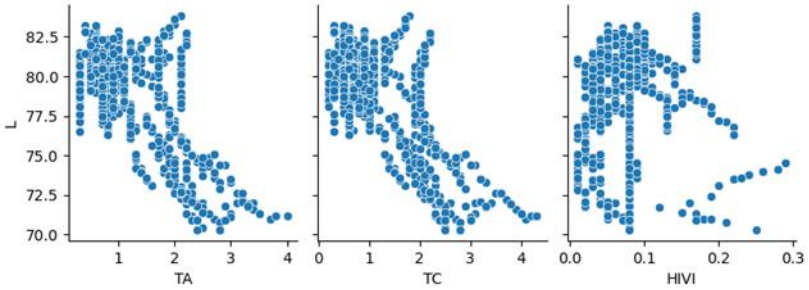
Variables TC, HB, and D were excluded from the model due to their high correlation with other independent variables, which could lead to multicollinearity issues. After implementing all modifications, the final empirical model for group 1 is a refined iteration of model (2):

$$\ln L_{it} = \ln GDP_{it}\beta_1 + \ln(\frac{1}{HIVI_{it}})\beta_2 + M_{it}\beta_3 + \ln(\sqrt{P_{it}})\beta_4 + TA_{it}\beta_5 + \ln BMI_{it}\beta_6 + \ln SY_{it}\beta_7 + GI_{it}\beta_8 + \ln\left(\frac{1}{AC_{it}}\right)\beta_9 + PP_{it}\beta_{10} + \alpha_i + \epsilon_{it} \quad (5)$$

For group 3 model, variables TA, TC, and $HIVI$ underwent an inverse transformation due to their inverse relationship with the dependent variable, as depicted in Figure 4.4.4.

FIGURE 4.4.4

DEVELOPED COUNTRIES PAIRPLOT (L, TA, TC, HIVI)

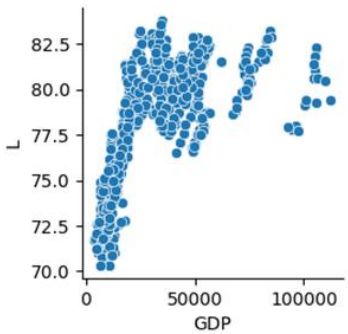


Source: Generated by author.

Furthermore, a logarithmic transformation is necessary for both the dependent variable L and the independent variable GDP to linearize the logarithmic relationship between these two variables as shown in Figure 4.4.5. BMI and SY underwent a square root transformation to better align their relationship with L as illustrated in Figure 4.4.6.

FIGURE 4.4.5

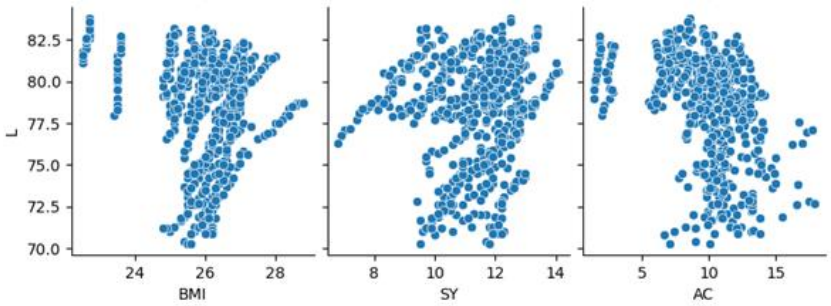
DEVELOPED COUNTRIES PARIPLOT (L, GDP)



Source: Generated by author.

FIGURE 4.4.6

DEVELOPED COUNTRIES PARIPLOT (L, BMI, SY)



Source: Generated by author.

Finally, variables TC, TA, $HIVI$, BMI, and SY underwent a logarithmic transformation to offset the logarithmic transformation of the dependent variable. Table 4.4.2 presents the correlation matrix of the 13 independent variables used in the model for group 3. Variables TA, TC, and D were excluded from the model to mitigate multicollinearity concerns. Observe that the variations in TA and TC can be accounted for by the independent variable BMI. Additionally, there are other independent variables to assess vaccine coverage rates. Therefore, omitting these variables will not compromise the reliability of the results. The final model for group 3 represents a refinement of model (2):

$$\ln L_{it} = \ln GDP_{it}\beta_1 + \ln\left(\frac{1}{HIVI_{it}}\right)\beta_2 + HB_{it}\beta_3 + M_{it}\beta_4 + P_{it}\beta_5 + \ln(\sqrt{BMI_{it}})\,\beta_6 + \ln\left(\sqrt{SY_{it}}\right)\beta_7 + GI_{it}\beta_8 + \ln\left(\sqrt{AC}_{it}\right)\beta_9 + PP_{it}\beta_{10} + \alpha_i + \epsilon_{it} \quad (6)$$

TABLE 4.4.2

CORRELATION MATRIX (GROUP 3)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) lnGDP	1.000												
(2) lnreTC	0.636	1.000											
(3) lnreTA	0.622	0.985	1.000										
(4) lnreHIVI	-0.183	-0.074	-0.029	1.000									
(5) lnsqrtBMI	-0.063	0.307	0.336	0.175	1.000								
(6) lnsqrtSY	0.308	-0.020	-0.032	0.079	0.208	1.000							
(7) PP	-0.083	0.156	0.113	-0.251	-0.043	-0.493	1.000						
(8) GI	0.144	0.134	0.143	0.308	-0.084	-0.085	0.001	1.000					
(9) lnsqrtAC	-0.217	0.005	0.010	0.135	0.223	0.002	0.084	0.081	1.000				
(10) HB	-0.247	-0.231	-0.234	0.204	0.113	-0.023	0.040	-0.053	-0.040	1.000			
(11) M	-0.293	-0.348	-0.366	0.223	0.042	0.276	-0.057	-0.101	-0.107	0.372	1.000		
(12) P	-0.124	-0.180	-0.189	-0.055	-0.176	-0.070	0.066	-0.023	0.013	0.123	0.351	1.000	
(13) D	-0.123	-0.191	-0.199	-0.065	-0.130	-0.043	0.082	-0.037	0.021	0.144	0.363	0.916	1.000

Note: lnGDP = ln GDP; lnreHIVI = ln($\frac{1}{HIVI}$); lnsqrtBMI = ln($\sqrt{BMI}$); lnsqrtSY = ln($\sqrt{SY}$).

Source: Generated by author.

The model for group 2 countries went through the same procedure of modification and variable selection. After modifications, the final model for group 2 is a refinement of model (4) that dropped HB, TA, TC, D and M because of their high correlations with other independent variables:

$$\ln L_{it} = \ln GDP_{it}\beta_1 + \ln\left(\frac{1}{HIVI_{it}}\right)\beta_2 + \ln\left(\sqrt{P}\right)\beta_3 + \ln BMI_{it}\beta_4 + \ln SY_{it}\beta_5 + GI_{it}\beta_6 + \ln AC_{it}\beta_7 + PP_{it}\beta_8 + \alpha_i + \xi_t + AG_{it} + ID_{it} + UFD_{it} + \epsilon_{it} \quad (7)$$

Similarly, correcting for multicollinearity of group 2 model will not affect the reliability of the results. After describing the data and the final models for three different groups of countries, the following section includes the results of the one-way FE OLS panel regressions and the analysis of the results.

Result

This analysis recognizes the potential impact of heteroskedasticity on the reliability of the final model estimates. To address this, this analysis employed the Modified Wald Test to identify groupwise heteroskedasticity within the fixed-effect regression models for each of the three country groups. Table 5.1.1 in Appendix C shows the results of the Modified Wald Test in three models (Baum, 2001). The results indicate the presence of groupwise heteroskedasticity across all models supported by the low probability values associated with the Chi-squared statistic.

To ensure the validity of the inferences drawn from the models, this analysis employs robust standard errors to adjust for the heteroskedasticity detected. Tables 5.1.2, 5.1.3, and 5.1.4 present the results of the robust FE-OLS regression analyses. Comparative evaluation of these three robust estimations yields significant insights into the determinants and potential improvements of life expectancy. Furthermore, it provides compelling evidence of the relevance and precision of the ET within the context of LDCs.

TABLE 5.1.2

ROBUST FE-OLS ESTIMATION RESULTS (GROUP 1)

lnL	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
lnGDP	.036	.02	1.82	.076	-.004	.076	*
lnBMI	1.426	.403	3.54	.001	.613	2.24	***
lnSY	.068	.04	1.72	.093	-.012	.148	*
lnreAC	-.003	.003	-0.89	.376	-.01	.004	
lnreHIVI	.031	.011	2.91	.006	.01	.052	***
lnsqrtP	.059	.036	1.62	.113	-.014	.132	
GI (Omitted)	0	.	.	.	.	.	
PP (Omitted)	0	.	.	.	.	.	
TA	-.001	.001	-0.93	.358	-.002	.001	
M	-.001	0	-2.21	.032	-.002	0	**
Constant	-.784	1.153	-0.68	.5	-3.109	1.54	

Mean dependent var	4.062	SD dependent var	0.122
R-squared	0.749	Number of obs	704
F-test	26.632	Prob > F	0.000
Akaike crit. (AIC)	-3043.155	Bayesian crit. (BIC)	-3006.701

*** $p<.01$, ** $p<.05$, * $p<.1$

Note: lnGDP = ln(GDP); lnBMI = ln(BMI); lnSY = ln SY; lnreAC = ln$(\frac{1}{AC})$; lnreHIVI = ln$(\frac{1}{HIVI})$; lnsqrtHB = ln $\sqrt{HB}$; lnsqrtP = ln $\sqrt{P}$; lnsqrtD = ln $\sqrt{D}$.

Source: Generated by author.

TABLE 5.1.3

ROBUST FE-OLS ESTIMATION RESULTS (GROUP 2)

lnL	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
lnsqrtP	.133	.04	3.35	.001	.054	.213	***
lnreHIVI	.015	.007	2.36	.02	.002	.029	**
lnAC	.002	.002	1.11	.268	-.001	.005	
lnSY	.059	.025	2.40	.018	.01	.109	**
lnBMI	.575	.146	3.93	0	.285	.866	***
lnGDP	.025	.01	2.52	.013	.005	.046	**
AG	.005	.005	0.86	.393	-.006	.015	
ID	.003	.005	0.61	.542	-.008	.014	
UFD	.014	.005	2.73	.007	.004	.024	***
GI (Omitted)	0	.	.	.	.	.	
PP (Omitted)	0	.	.	.	.	.	
Constant	1.707	.402	4.24	0	.909	2.505	***

Mean dependent var	4.239	SD dependent var	0.114
R-squared	0.497	Number of obs	1552
F-test	.	Prob > F	.
Akaike crit. (AIC)	-7588.679	Bayesian crit. (BIC)	-7545.900

*** $p<.01$, ** $p<.05$, * $p<.1$

Note: lnsqrtP = ln$(\sqrt{P})$; lnreHIVI = ln$(\frac{1}{HIVI})$; lnAC = ln AC; lnSY = ln SY; lnBMI = ln BMI; lnGDP = ln GDP.

Source: Generated by author.

TABLE 5.1.4

ROBUST FE-OLS ESTIMATION RESULTS (GROUP 3)

lnL	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
lnGDP	.009	.011	0.84	.408	-.013	.032	
lnreHIVI	.005	.003	1.38	.175	-.002	.011	
lnsqrtBMI	1.63	.348	4.68	0	.924	2.335	***
lnsqrtSY	.207	.034	6.04	0	.138	.276	***
GI (Omitted)	0	.	.	.	.	.	
PP (Omitted)	0	.	.	.	.	.	
AC	-.001	.002	-0.57	.571	-.005	.003	
HB	0	0	2.11	.042	0	0	**
M	0	0	-1.31	.197	0	0	
P	0	0	1.42	.165	0	0	
Constant	1.339	.499	2.68	.011	.327	2.351	**

Mean dependent var	4.363	SD dependent var	0.041
R-squared	0.787	Number of obs	608
F-test	52.183	Prob > F	0.000
Akaike crit. (AIC)	-4232.949	Bayesian crit. (BIC)	-4197.667

*** $p<.01$, ** $p<.05$, * $p<.1$

Note: lnGDP = ln GDP; lnreHIVI = ln($\frac{1}{HIVI}$); lnsqrtBMI = ln($\sqrt{BMI}$); lnsqrtSY = ln($\sqrt{SY}$).

Source: Generated by author.

According to the three result tables, the high R-square values of group 1 (around 0.75) and group 3 (around 0.79) models indicate that the independent variables included are significant predicators of average life expectancy changes within the same entity over different time period. Conversely, the R-square value of group 2 model is relatively low (around 0.5), which indicates that the model did not effectively explain the variation in the average life expectancy changes in developing countries. The wide disparities among developing countries (LDCs excluded) in epidemiological, economic, technological, and educational domains lead to significant heterogeneities. It is difficult to thoroughly account for all different heterogeneities within this category.

According to the results of Table 5.1.2, the factors most significantly affecting average life expectancy in LDCs are HIV incidence and BMI with p-values less than

0.01. In developing countries (excluding LDCs), the most critical factors are BMI, Polio vaccine coverage rates, and the binary variable indicating whether infant mortality rates are below the average level of developed countries. However, in developed countries, the factors that most significantly impact average life expectancy are BMI and years of schooling. It is noteworthy that among the most influential factors, both LDCs and developing countries (excluding LDCs) include variables directly related to infectious diseases, such as HIV incidence and Polio vaccine coverage rates. However, in Table 5.1.4, no variable related to infectious diseases shows a significant impact on average life expectancy in the models for developed countries.

For instance, HIV incidence is negatively correlated with average life expectancy in both LDCs and developing countries (excluding LDCs), with a more substantial negative impact in LDCs than in developing countries (excluding LDCs). Conversely, HIV incidence in developed countries shows statistically insignificant effects with minimal coefficients. These results indicate the accuracy of the ET, suggesting that LDCs, being in the overlapping status of stages 1 and 2 of the epidemiological transitions, are more susceptible to infectious diseases, even though Omran did not witness the prevalence of HIV and the emergence of the LDC category. This study reveals that HIV, as a currently incurable epidemic with a significant global patient population, has the most adverse effect on average life expectancy in LDCs, while its impact is much lower in developed countries due to the completeness of the epidemiological transition.

Similarly, the study found that Polio vaccine coverage rates have a statistically significant positive effect on enhancing average life expectancy in developing countries. In LDCs, however, there was no statistically significant positive impact of

increased Polio vaccine coverage rates on average life expectancy. This finding further corroborates the ET perspective that infectious diseases are challenging to control in LDCs because the sanitation transition, as part of the epidemiological transition, did not occur in LDCs, which limits the effectiveness of medical technology such as vaccination in these countries.

The findings of this study underscore the importance for policymakers in LDCs to recognize that controlling infectious diseases, with HIV at the forefront, remains a key strategy for enhancing the average life expectancy in modern world. Moreover, these findings highlight a significant challenge to resolving this issue: medical technologies, such as vaccination, are unable to achieve their potential impact in LDCs where the epidemiological transition did not complete. The study suggests that LDCs should focus on promoting public awareness and education regarding the prevention of infectious diseases. Additionally, this analysis suggests LDCs to accelerate the process of epidemiological transition by increasing per capita income, investing in nutrition and food, ensuring the availability of basic education, and establishing accessible public health infrastructures.

For other developing countries that already possess relatively advanced medical system, this research suggests that the widespread adoption of medical technologies, such as vaccination, plays a crucial role in positively influencing the average life expectancy of their populations. Furthermore, the negative impact of infectious diseases on average life expectancy, such as HIV, is significantly reduced in developing countries comparing with LDCs. However, this study is unable to offer sufficient epidemiology-related advice for current developed countries, as the independent variable directly associated with degenerative diseases, alcohol consumption, was statistically insignificant in the final model. Moreover, independent

variables with explanatory power were unrelated to diseases in group 3 model. The following section includes more discoveries and limitations of this analysis.

Discussion

This section aims to provide guidance for the appropriate and efficient use of the findings from this analysis. This section highlights several additional discoveries through the research that are not focused on in detail. Firstly, this analysis assumed in the model section that time-related heterogeneity, when all countries could not transition across epidemiological stages within the timeframe, would cease to exist. This analysis made this assumption because controlling for time-related heterogeneity would require an additional sixteen variables, compromising degrees of freedom and potentially causing severe heteroskedasticity in a sample size that is not sufficiently large. While the assumption does not affect the validity of the results, considering time-related heterogeneity in an appropriate manner could undoubtedly enhance the model's accuracy and explanatory power. The study calls for future scholars to find a balance between unobserved heterogeneity and the risk of overfitting the model. The second discovery raised concerns to the availability of data. As mentioned earlier, data on average life expectancy and related variables, particularly before the year 2000, are often not available for Least Developed Countries (LDCs) and are mostly in annual reports, which greatly limits the sample size, especially in comparison analyses such as this study, where observations for individual models are further reduced.

The third discovery is related to the selection of variables. Among the three proxy variables used in this study, only the variable assessing the level of under-five children's deaths (UFD) showed statistical significance in the final results. This analysis encourages future researchers to explore other proxy variables to better assess the epidemiological stages of developing countries. Additionally, this analysis did not include the Gini Index and Political Philosophy as independent variables in

the final models due to multicollinearity. Future studies should explore alternative variables to assess inequality and the natural environment. Lastly, the study did not find any statistically significant correlation between infectious disease vaccine coverage rates and average life expectancy in developed countries. This analysis calls upon future researchers to examine and explain this phenomenon.

In conclusion, this analysis revisits the ET, a theory proposed in the 1960s, to highlight its practicality against contemporary challenges in LDCs and calls for a renewed examination of the theory's importance. As stated by Omran, health is a dependent variable of epidemiology, not vice-versa (Omran, 1998). This analysis provides literature support for improving life expectancy in LDCs from epidemiological perspective and, through comparative analysis, offers new research directions towards the reduction of global life expectancy disparities.

APPENDIX A

TABLE 4.1.1

GINI INDEX SOURCES

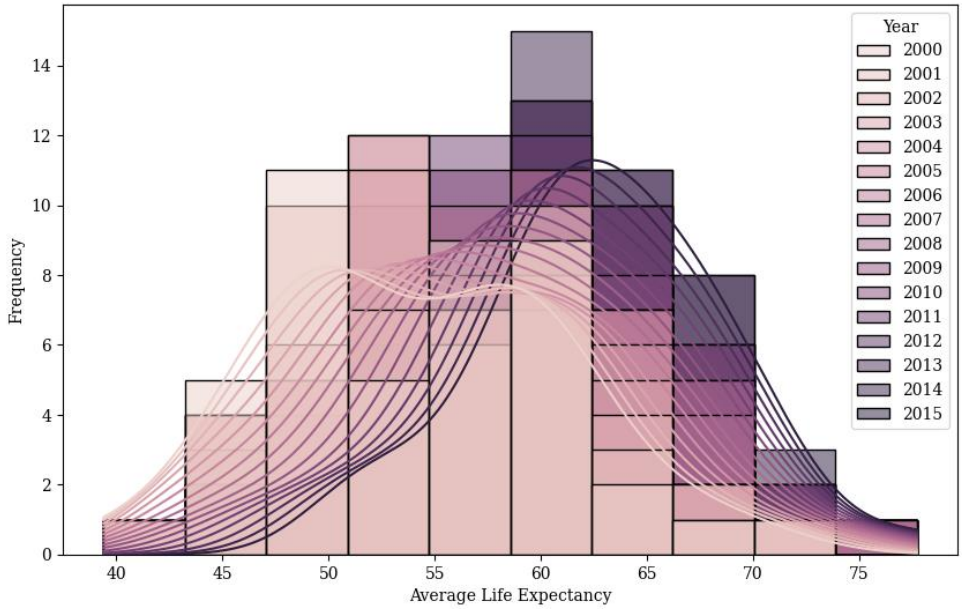
Source	**Country Name**
The Federal Research Economic Data (FRED)	Djibouti (2013); Kiribati (2019); Sao Tome and Principe (2010); Solomen Islands (2012); Timor-Leste (2014); Czechia (2019); Belize (1999); Cape Verde (2015); Fiji (2019); Iran (2019); Iraq (2012); Cape Verde (2015); Micronesia (2013); Kyrgyz Republic (2019)
The World Bank	Maldives (2019); Montenegro (2018); Samoa (2013); Seychelle (2018); St. Lucia (2016); Syrian Arab Republic (2003); Tonga (2015); Trinidad and Tobago (1992); Vanuatu (2019); Venezuela (2006)
Statista	Cuba (2024); Brunei Darussalam (2024)
The United Nations	Barbados (2017-2021)
Confidus Solutions	Eritrea (unknown)
World Economics	Other countries excluding Antigua and Barbuda, Somalia, Grenada, and St. Vincent and the Grenadines (2019)

Note: The number in the parentheses represents the corresponding year of the Gini index used in this analysis. The references section of this analysis includes details of these 6 sources.

Source: Generated by author.

FIGURE 4.3.2

AVERAGE LIFE EXPECTANCY DISTRIBUTION (GROUP 1)

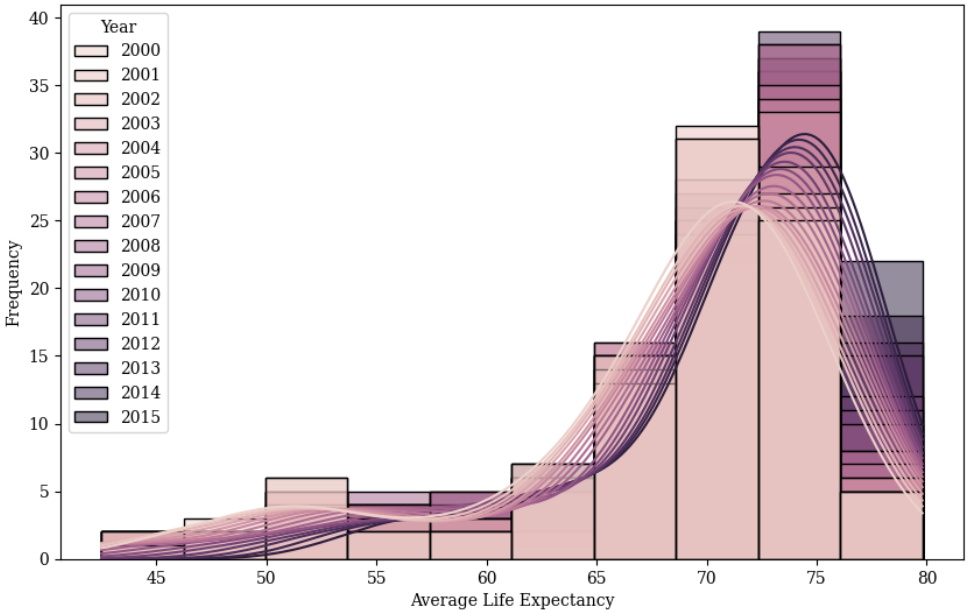


Source: Generated by author.

APPENDIX B

FIGURE 4.3.3

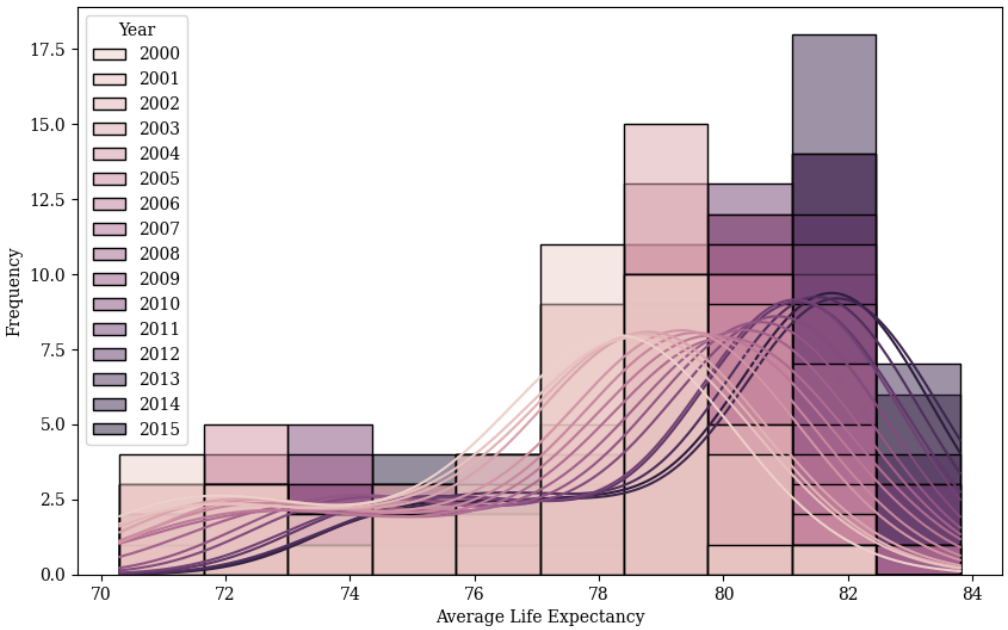
AVERAGE LIFE EXPECTANCY DISTRIBUTION (GROUP 2)



Source: Generated by author.

FIGURE 4.3.4

AVERAGE LIFE EXPECTANCY DISTRIBUTION (GROUP 3)



Source: Generated by author.

APPENDIX C

TABLE 5.1.1

MODIFIED WALD TEST

$H_0: \sigma_i^2 = \sigma^2, \forall i$	**Group 1**	**Group 2**	**Group 3**
Chi-squared Statistics	48806.52	4.8e+05	4308.90
Prob > χ^2	0.0000	0.0000	0.0000

Source: Generated by author.

Reference

1. The United Nations. (2015). The Sustainable Development Agenda. Retrieved from https://www.un.org/sustainabledevelopment/development-agenda/.
2. Leboeuf, C. (2018). Global indicator framework for the Sustainable Development Goals
3. Halicioglu, F. (2011). Modeling life expectancy in Turkey. Economic modelling, 28 (5), 2075-2082.
4. Liou, L., Joe, W., Kumar, A., & Subramanian, S. V. (2020). Inequalities in life expectancy: An analysis of 201 countries, 1950–2015. Social Science & Medicine, 253, 112964.
5. Alwan, A., & World Health Organization. (2016). Five years in action: strengthening public health in the Region and beyond.
6. 5th United Nations Conference on the Least Developed Countries (LDC5). (2023). Profiles of LDCs. Retrieved from https://www.un.org/ldc5/content/profiles-ldcs.
7. Novak, A., Čepar, Ž., & Trunk, A. (2016). The role of expected years of schooling among life expectancy determinants. International Journal of Innovation and Learning, 20(1), 85-99.
8. Sen, A. (1991). What did you learn in the world today?. American Behavioral Scientist, 34(5), 530-548.
9. Preston, S. H. (1975). The changing relation between mortality and level of economic development. Population studies, 231-248.
10. Gilligan, A. M., & Skrepnek, G. H. (2015). Determinants of life expectancy in the Eastern Mediterranean Region. Health Policy and Planning, 30(5), 624-637.
11. Anatoly, V. (2020). The epidemiological transition and its interpretations. Демографическое обозрение, 7(5), 4-41.
12. Omran, A. R. (1971). The Epidemiologic Transition. Milbank Memorial Fund Quarterly, 49(1), 509-538.
13. Omran, A. R. (1998). The epidemiologic transition theory revisited thirty years later. World health statistics quarterly, 53(2, 3, 4), 99-119.
14. Defo, B. K. (2014). Demographic, epidemiological, and health transitions: are they relevant to population health patterns in Africa?. Global health action, 7(1), 22443.
15. Nnebue, C. C. (2010). The epidemiological transition: policy and planning implications for developing countries. Nigerian journal of medicine, 19(3), 250-256.
16. Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2013). Life expectancy. Our world in data.
17. Oeppen, J., & Vaupel, J. W. (2002). Broken limits to life expectancy. Science, 296(5570), 1029-1031.

18. Burger, O., Baudisch, A., & Vaupel, J. W. (2012). Human mortality improvement in evolutionary context. Proceedings of the National Academy of Sciences, 109(44), 18210-18214.

19. Olshansky, S. J., & Ault, A. B. (1986). The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. The Milbank Quarterly, 355-391.

20. Fialho, D. (2012). Altruism but not quite: the genesis of the Least Developed Country (LDC) category. Third World Quarterly, 33(5), 751-768.

21. Livi Bachchi, M. (2010). Demograficheskaya istoriya Evropy [The demographic history of Europe]. Saint Petersburg: Alexandria.

22. Boutayeb, A. (2010). The burden of communicable and non-communicable diseases in developing countries. Handbook of disease burdens and quality of life measures, 531.

23. Frenk, J., & Gomez-Dantes, O. (2011). The triple burden: disease in developing nations. Harvard International Review, 33(3), 36.

24. The World Health Organization. (n.d.-a). Global health estimates: Leading causes of death. Retrieved from: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death.

25. The Centers for Disease Control and Prevention. (2022). About HIV. Retrieved from: https://www.cdc.gov/hiv/basics/whatishiv.html.

26. Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., ... & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science, 220(4599), 868-871.

27. Hsu, D. C., & O'Connell, R. J. (2017). Progress in HIV vaccine development. Human vaccines & immunotherapeutics, 13(5), 1018-1030.

28. Lu, D. Y., Lu, T. R., Zhu, H., Yarla, N. S., Che, J. Y., & Wu, H. Y. (2017). Pathogenesis studies of HIV/AID, a general viral topic. EC Orthopaedics, 5(4), 138.

29. Burton, D. R., Ahmed, R., Barouch, D. H., Butera, S. T., Crotty, S., Godzik, A., ... & Wyatt, R. (2012). A blueprint for HIV vaccine discovery. Cell host & microbe, 12(4), 396-407.

30. Deeks, S. G., Lewin, S. R., Ross, A. L., Ananworanich, J., Benkirane, M., Cannon, P., ... & Zack, J. (2016). International AIDS Society global scientific strategy: towards an HIV cure 2016. Nature medicine, 22(8), 839-850.

31. The World Health Organization. (n.d.-b). Essential Programme on Immunization. Retrieved from: https://www.who.int/teams/immunization-vaccines-and-biologicals/essential-programme-on-immunization.

32. Felice, E., Andreu, J. P., & D'Ippoliti, C. (2016). GDP and life expectancy in Italy and Spain over the long run: A time-series approach. Demographic research, 35, 813-866.

33. Wikipedia. (2023). Body Mass Index. Retrieved from: https://en.wikipedia.org/wiki/Body_mass_index.

34. Walls, H. L., Backholer, K., Proietto, J., & McNeil, J. J. (2012). Obesity and trends in life expectancy. Journal of obesity, 2012; De Backer, G., De Bacquer, D., & Prospective Studies Collaboration. (2009). Body-mass index and cause-specific mortality in 900.000 adults: collaborative analyses of 57 prospective studies. Lancet.

35. Peeters, A., Barendregt, J. J., Willekens, F., Mackenbach, J. P., Mamun, A. A., Bonneux, L., & for NEDCOM, the Netherlands Epidemiology and Demography Compression of Morbidity Research Group*. (2003). Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Annals of internal medicine, 138(1), 24-32.

36. Di Angelantonio, E., Bhupathiraju, S. N., Wormser, D., Gao, P., Kaptoge, S., De Gonzalez, A. B., ... & Hu, F. B. (2016). Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. The Lancet, 388(10046), 776-786.

37. Mushtaq, M. U., Gull, S., Khurshid, U., Shahid, U., Shad, M. A., & Siddiqui, A. M. (2011). Prevalence and socio-demographic correlates of stunting and thinness among Pakistani primary school children. BMC public health, 11(1), 1-12.

38. Mijinyawa, M. S., Yusuf, S. M., Gezawa, I. D., Musa, B. M., & Uloko, A. E. (2014). Prevalence of thinness among adolescents in Kano, Northwestern Nigeria. Nigerian Journal of Basic and Clinical Sciences, 11(1), 24-29.

39. Corrao, G., Bagnardi, V., Zambon, A., & La Vecchia, C. (2004). A meta-analysis of alcohol consumption and the risk of 15 diseases. Preventive medicine, 38(5), 613-619.

40. The World Bank. (2024). GiniIndex. Retrieved from: https://databank.worldbank.org/metadataglossary/gender-statistics/series/SI.POV.GINI.

41. Te, H. S., & Jensen, D. M. (2010). Epidemiology of hepatitis B and C viruses: a global overview. Clinics in liver disease, 14(1), 1-21.

42. Alavian, S. M. (2011). New globally faces of hepatitis B and C in the world. Gastroenterology and Hepatology from Bed to Bench, 4(4), 171.

43. Assaad, F. (1983). Measles: summary of worldwide impact. Reviews of infectious diseases, 5(3), 452-459.

44. Mehndiratta, M. M., Mehndiratta, P., & Pande, R. (2014). Poliomyelitis: historical facts, epidemiology, and current challenges in eradication. The Neurohospitalist, 4(4), 223-229.

45. Sharma, N. C., Efstratiou, A., Mokrousov, I., Mutreja, A., Das, B., & Ramamurthy, T. (2019). Diphtheria. Nature Reviews Disease Primers, 5(1), 1-18.

46. Xiao, Y., Li, H., & Fleisher, B. M. (2015). The earnings effects of health and health-related activities: a panel data approach. Applied Economics, 47(14), 1407-1423.

47. Wikipedia. (2024). List of socialist states. Retrieved from: https://en.wikipedia.org/wiki/List_of_socialist_states.

48. Federal Reserve Economic Data. (n.d.). Gini Index. Retrieved from: https://fred.stlouisfed.org/searchresults?st=Gini+Index.

49. The World Bank. (n.d.). Gini Index. Retrieved from: https://data.worldbank.org/indicator/SI.POV.GINI.

50. Statista. (n.d.). Economic Inequality – Worldwide. Retrieved from: https://www.statista.com/outlook/co/socioeconomic-indicators/economic-inequality/worldwide.

51. United Nations. (2016). Sub-regional country programme document for Barbados and the Organization of Eastern Caribbean States (2017-2021). Retrieved from: https://www.undp.org/sites/g/files/zskgke326/files/migration/bb/undp_bb_CPD-BRB-2017-01.pdf.

52. World Economics. (n.d.). Gini Coefficient. Retrieved from: https://www.worldeconomics.com/Indicator-Data/Inequality/Gini-Coefficient.aspx.

53. Kannisto, V. (2001). Mode and dispersion of the length of life. Population: An English Selection, 159-171.

54. Kovchegov, Y. (2022). A new life of Pearson's skewness. Journal of Theoretical Probability, 35(4), 2896-2915.

55. Baum, C. F. (2001). Residual diagnostics for cross-section time series regression models. The Stata Journal, 1(1), 101-104.