Health and Income: A Dynamic Panel Data Model

Azad, Kalam

University of Ottawa

15 April 2020

Health and Income: A Dynamic Panel Data Model

Kalam Azad*

September 6, 2020

Abstract

A panel data approach has been used to find the effect of life expectancy on GDP per capita for 182 countries over 1960-2015. Panel data models are considered and compared their estimates. Accounting for country fixed effects and dynamics in GDP per capita, our models document a significantly positive and robust effect of life expectancy on GDP per capita. The dynamic fixed effects model provides that GDP per capita increases by 3.5% with a 1% rise in life expectancy: On average around a 20-year rises in life expectancy provides about $10,913 more mean GDP per capita which is large in magnitude. Results remain very similar when using instrumental variable model. Our investigation suggests that an improvement in life expectancy increases GDP per capita by reducing fertility and increasing saving.

Keywords: GDP Per Capita, Life Expectancy, Mortality, Survival, Saving, Fertility

JEL-Classification: I15

*Azad: University of Ottawa, Ottawa, ON, postal code: K1N6N5, Canada, email: kazad044@uottawa.ca. phone:613-710-8486. Acknowledgments: An anonymous referee, Jason Garred, Marcel Voia, Luc Patry and iHEA Conference participants for commenting on earlier version of this paper.

1 Introduction

Global health is in two regimes: unhealthy nations with high mortality rates and healthy nations with few child deaths. Healthier countries are more productive and thus contribute more in GDP per capita.[1] Economic growth in healthier countries are enough to provide public health care services because health as a byproduct automatically appear from economic growth.[2] Unfortunately, it is not true for unhealthier nations because their economic growth is a far away from enough to maintain health care for all citizens (Sachs, 2002). In addition, health with communicable and preventive diseases in these countries reduce economic growth and consequently deteriorate health conditions. Health in these countries can be improved from their own economic growth and a large philanthropic resources from developed countries (Ashraf et al., 2008). As described the whole health coverage may not be possible, a combination of both can be effective to ameliorate health in poor nations. To reduce global health inequality, an integrated efforts of global community from humanitarian aspects has commenced in 2000 with a set of goals called the Millennium Development Goals (MDGs): It achieved considerable success in reducing mortality even if the target level of two-thirds reduction of child deaths has not reached by 2015 and thus increases life expectancy at birth.[3] Following the success from

[1] We use health to refer to life expectancy at birth or general health outcome or other health outcomes such as mortality rates and survival rate throughout the paper.

[2] We also use GDP and income interchangeably to refer to GDP per capita in this paper.

[3] MDGs are 8 goals: to alleviate extreme poverty and hunger; to achieve universal primary education; to promote gender equality and empower women; to reduce child mortality; to improve maternal health; to prevent HIV/AIDS, malaria and other preventive and communicable diseases; to make sure environmental sustainability and to develop a global partnership for development.

MDGs, the sustainable development goals (SDGs) set target of reducing child deaths under five to 25 per 1,000 by 2035.[4] These global attempts promote global health and economic growth and development. For example, WHO's commission reports, chaired by Jeffrey Sachs, on Macroeconomics and Health (2002):

"For individuals and families, health brings the capacity for personal development and economic security in the future. Health is the basis for job productivity, the capacity to learn in school, and the capability to grow intellectually, physically, and emotionally. As with the economic well-being of individual households, good population health is a critical input into poverty reduction, economic growth, and long-term economic development at the scale of whole societies".

"Health is a resource that enables every person to realize his or her potential and to contribute to the overall development of society"(WHO and Others, 2017).

Theoretical and empirical analyzes in growth and other economic fields provide evidence that health is a matter for economic growth. Health can affect economic growth through mechanisms: First, it affects per capita labor through fertility. Second, it scales saving up and hence capital formation. The existing literature has acknowledged the importance of channels via which health affects the GDP per capita. For instance, Aghion et al. (2010) and Barro (1996) underscore the significance of investigation of the channels through which health acts on growth: An improved life expectancy reduces mortality

[4]Mortality reducing to 25 per 1,000 is one of 17 goals in SDGs.

and so does fertility which in turn increases the economic growth. A higher life expectancy can save more and thus economic growth goes up (Zhang and Zhang, 2005). People with a longer life expectancy can reduce fertility which is based on the condition that they live in a low infant mortality environment (Aghion et al., 2010). So, both population and fertility go down through demographic transition with varying lags. This can lead to having more capital per labor to work with and hence productivity and growth rises. Thus, mechanisms are important to find how health works on economic growth.

There are a few papers that previously studied the direct link between general health outcome, measured by life expectancy at birth and GDP per capita on a big panel dataset over a long time period. To our knowledge, the first empirical research in this literature started from Barro (1996). He finds a significantly positive relationship between life expectancy and economic growth from 1960 to 1990 while Acemoglu and Johnson (2007) find no impacts of life expectancy on GDP per capita over 1940-1980.[5] Thus, there is a mixed in empirical evidence. However, these papers do not consider time regimes from 2000 to 2015 in which Millennium Development Goals (MDGs) happened when health has substantially improved by reducing child mortality rates globally. Furthermore, time regimes 1980 and onwards include global public health goods such as AIDS/HIV and AMR which did not occur before 1980. The global mortality instruments constructed byAcemoglu and Johnson (2007) may not reflect mortality caused by HIV/AIDS as their instruments are based on diseases occurred in 1940. They also suggest that their results may not apply for present

[5]Their main time regimes are 1940-1980, however they extended to 2000 to check robustness.

time because of these global public health goods which were not occurred at the time regimes considered in their research (Acemoglu and Johnson, 2007, p. 4). In consequence, these time regimes are important for establishing the link between income and health. Hence, we enlarge the dimensions of our panel data unlikely to Acemoglu and Johnson (2007) or Barro (1996): Both cross-section dimension, N and time series dimension, T increases; N goes to 182 countries and T to 56 years. This allows us find consistent estimates across cross-section units given stationary process of panel units.

A number of papers related to ours-using different health indicators rather than life expectancy-investigates the connection between income and health. For example, Ehrlich and Lui (1991) and Bhargava et al. (2001) discover a positive impacts of improved health on income by measuring health as probabilities of survival from ages 0 to 25 and 50 to 75, adult survival rates over 1960-2010 respectively while Aghion et al. (2010) and Meltzer (1992) find that there is negative impact of mortality rates on income.

Our interest in this paper is to find a relationship between health and GDP per capita using panel data estimators. To attain this relation, this paper follows several strategies: First, we use a dynamic panel model which captures country fixed effects and persistence in GDP per capita (or dynamics of GDP per capita). The inclusion of persistence in our model gives us different results relative to a static panel model. Second, since health is an endogenous variable, without addressing this issue our estimates may be biased and inconsistent. To overcome this problem, we follow an instrumental variable (IV) method. Using lags as internal instruments, we employ the generalized method of moments

(GMM) estimator. Third, using both estimators, we undertake several strategies to check robustness of our results such as alternative health measures, time regimes and outliers. Finally, our paper examines channels through which health acts on income per capita.

In the literature, there are theoretical analyzes concerning the dynamic relationship between income and health. For example, Malthus postulates a negative dynamic relationship between mortality and economic growth (Preston, 1975). Also, studies from Acemoglu et al. (2019), Cuestas and Garratt (2011), and Chang et al. (2006) demonstrate that there is a high persistence in GDP. Thus, our conceptual framework of dynamic panel model follows from this. Controlling the dynamics of GDP, we can account for factors affected past also affect on current GDP. We can estimate parameters consistently using the overall amounts of persistence in GDP by employing the dynamic fixed effects (DFE) and GMM models which allow us to find the long-run effects of health. As a result, our conceptual framework and data differ from others: This analysis provides different results from others.

Our contributions in this paper are: First, we find a significantly positive and robust effect of life expectancy at birth on GDP per capita. Second, we use an updated yearly dataset from 1960 to 2015 for 182 countries which covers the entire world. Third, our analysis discovers channels via which life expectancy affects on GDP per capita. Our preferred estimates from the dynamic within estimator-baseline model-indicates that GDP per capita rises by around 3.5% due to a 1% rise in life expectancy at birth in the long-run : On average, if life expectancy rises by 20 years (for example, from 1950 to

1970) then mean GDP per capita rises around $10,913 more which is a very large magnitude.[6] We obtain a very similar estimate from employing GMM estimator using internal instruments. Our analysis shows that health raises income by decreasing fertility and increasing saving.

The paper is organized as follows: Section 2 denotes the data sources and description. Section 3 explains the econometric models and empirical results. Section 4 provides the mechanisms and test of mechanisms and concluding remarks are given in section 5.

2 Data Sources and Description

In our analysis, GDP per capita (constant 2010 U.S. dollars) is our outcome and life expectancy at birth are independent variables: Data for both variables are used from the World Bank's World Development Indicators (World Bank, 2018). Other health indicators we use from this data source include under-five child mortality (per 1,000 live births), maternal mortality (per 1,000 live births), male adult mortality (per 1,000 male adult), male survivals to age 65 (percentage of cohort), saving (percentage of GDP) and total fertility rate (the number of children born per woman). Table 1 presents summary statistics of outcome and regressors in our model.

[6]Dynamic within or dynamic fixed effects or within estimators are interchangeably used in this paper.

Table 1: Descriptive Statistics

	Obs	Mean	S.D.
GDP per capita	8,097	$9,743.86	$14,985.53
Life expectancy at birth	10,088	63.18	11.43
Child mortality under five per 1,000 births	8,960	80.65	78.99
Maternal mortality per 1,000 births	4,654	273.61	364.85
Male adult mortality per 1,000	10,013	274.17	115.37
Male survival rate to age 65	10,024	58.45	16.08
Saving	8,097	$59,246,602,290.23	$256,132,804,227.09
Fertility	10,092	4.158	2.042

Note: The detailed description of variables in the text and sources from where they are used.

3 Panel Data Estimators

Our paper uses two panel estimators to model the relationship between health and income: Dynamic fixed effects (dynamic within), and Arellano-Bond GMM (dynamic IV) estimators.

3.1 Dynamic Fixed Effects Model (Baseline)

Using dynamic panel models, this paper constructs a relationship between GDP per capita and health. Our dynamic within model is:

$$y_{it} = \beta h_{it} + \sum_{l=1}^{q} \delta_l y_{i(t-l)} + \alpha_i + \gamma_t + u_{it} \tag{1}$$

where, $i = 1, 2, \ldots, 182$, indicates country over the years $t = 1960, 1961, \ldots, 2015$ and $l = 1, 2, \ldots, q$ are lags. In this model y, h, y_{t-l} and u represent log of GDP per capita, log of health, lagged dependent variables and the error term respectively. Here, health indicates life expectancy at birth and alternative measures of health (Mortality rates and male survivals). Our interest lies in estimating β: It is expected to have a positive value indicating a higher GDP is

correlated with an improvement in life expectancy. δ_l captures the persistence in the outcome variable which may be due to the effects of policies that have previously been taken in each country. α_i and γ_t capture the country and year fixed effects in our model respectively. The u contains all other factors excluded from the model.

Our dynamic panel within estimator is consistent provided GDP per capita is stationary and health is exogenous. The identification of β in model (1) depends on zero conditional mean of error u_{it}:

Assumption 1: $\mathbf{E}(u_{it} \mid h_{it}, y_{i(t-1)}, \ldots, y_{i(t-q)}, \alpha_i, \gamma_t) = 0$, for all h_{it}, $y_{i(t-1)}$, $\ldots, y_{i(t-q)}$, α_i and γ_t.

This is a fundamental assumption conveying conditional mean of u_{it} given the values of health, past GDP, country and year fixed effects. The implication of this assumption is that error terms are conditionally uncorrelated: health, lagged GDP, country and year fixed effects are exogenous.

To find the long-run effects of health on income, we derive the following formula accounting for the persistence in income. If outcome and explanatory variables are persistent, then in the steady-state, $y_{it} = y_{ss}$, $h_{it} = h_{ss}$. For simplicity, ignoring the country and year fixed effects as well as the error term in model 1, we can attain, $y_{it} = y_{it-l} = y_{ss}$ in steady-state. Using these in model 1, we obtain

$$y_{ss} = \beta h_{ss} + \sum_{l=1}^{q} \delta_l y_{ss}$$

$$\text{so, } y_{ss} = \frac{\beta h_{ss}}{1 - \sum_{l=1}^{q} \delta_l}$$

where, y_{ss} is a steady-state level of income in the long-run. The long-run cumulative effects of health is $\frac{\beta}{1 - \sum_{l=1}^{q} \delta_l}$, where l are lags, $1 \leq l \leq q$, $\sum_{l=1}^{q} \delta_l$ represents the sum of autoregressive coefficients indicating the overall amounts of persistence in income and $\sum_{l=1}^{q} \delta_l$ converges to $m \in (0, 1)$. Because the long-run effects of health are obtained after estimation, we use the estimated coefficients of β and δ. Therefore, the long-run formula becomes $\frac{\hat{\beta}}{1 - \sum_{l=1}^{q} \hat{\delta}_l}$, where $\sum_{l=1}^{q} \hat{\delta}_l$ converges to $m \in (0, 1)$ which implies income is stationary.

We can consistently estimate the dynamic effects of health controlling for the various lagged dependent variables. Table 2 presents the impacts of life expectancy on income. We control for one lag of outcome variable in column 1, Panel A. The estimated and long-run effects of life expectancy are positive and statistically significant: The estimated coefficient for life expectancy is 0.202 (s.e.= 0.091) while in the long-run it is 1.934 (s.e.= 0.677); a 1% rise in life expectancy contributes around 2% rise in GDP in the long-run. The persistence in income is 0.896 (s.e.= 0.073) which is less than 1: It implies income satisfies stationary process. Our estimate of β is consistent provided model 1 satisfies assumption 1 and GDP is stationary.[7][8]

[7]To find the stationary of GDP per capita, we conduct Fisher-type panel unit root test for our unbalanced panel data. The p-value of this test rejects the presence of unit root in GDP per capita. In Fisher-type unit root test, H_0: All panels contain unit root; H_1: At least one panel is stationary.

[8]The maximum lags we considered in our analysis is 5. To select preferred lags, we set the null hypothesis $H_0 : \gamma_{il} = 0$ on the following augmented Dickey-Fuller regression,

When we add one more lag in column 2, the estimated effect of life expectancy is around 0.221% (s.e.= 0.081%) while the estimate for cumulative long-run effects is around 3.162% (s.e.= 0.903%). Both estimates are higher than those of column 1. The persistence in GDP is 0.930 (s.e.= 0.041) which shows that GDP is stationary. With three lags of the outcome variable, our estimates in column 3 indicate that the amount of long-run effect of health is 3.511% with a standard error of 0.917%. This is our preferred specification. The sum of autoregressive coefficients is less than one indicating stationarity of GDP. Columns 4 and 5 with 4 and 5 lags of GDP respectively demonstrate very similar effects with the estimates of column 3.

Table 2: The effect of life expectancy on GDP per capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Dynamic within estimates					GMM estimates			
Health effect	0.202**	0.221***	0.095***	0.091***	0.095***	0.464**	0.471**	0.146***	0.138***	0.139***
	(0.091)	(0.081)	(0.016)	(0.015)	(0.015)	(0.247)	(0.200)	(0.024)	(0.024)	(0.025)
GDP per capita	0.896***	0.572***	1.070***	1.163***	1.142***	0.802***	0.492***	1.034***	1.126***	1.111***
first lag	(0.076)	(0.115)	(0.052)	(0.067)	(0.066)	(0.133)	(0.118)	(0.048)	(0.072)	(0.070)
GDP per capita		0.358***	-0.030	-0.136*	-0.080		0.347***	-0.021	-0.117	-0.073
second lag		(0.076)	(0.023)	(0.076)	(0.073)		(0.031)	(0.022)	(0.074)	(0.072)
GDP per capita			-0.067**	-0.022**	-0.064**			-0.050**	-0.018***	-0.055**
third lag			(0.030)	(0.006)	(0.028)			(0.026)	(0.007)	(0.025)
GDP per capita				-0.036**	-0.021**			-0.050**	-0.018***	-0.055**
fourth lag				(0.012)	(0.010)				(0.015)	(0.010)
GDP per capita					-0.012					-0.006
fifth lag					(0.010)					(0.008)
Long run	1.934***	3.162***	3.511***	3.000***	2.739***	2.348***	2.936***	3.954***	3.673***	3.349***
effect of health	(0.677)	(0.903)	(0.917)	(0.704)	(0.658)	(0.466)	(0.548)	(0.679)	(0.582)	(0.549)
Persistence in	0.896***	0.930***	0.973***	0.970***	0.966***	0.802***	0.840***	0.963***	0.962***	0.959***
GDP per capita	(0.076)	(0.041)	(0.005)	(0.005)	(0.006)	(0.133)	(0.090)	(0.007)	(0.007)	(0.008)
Panel unit root test										
of GDP (p-value)	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]					
AR2 test (p-value)						[0.652]	[0.000]	[0.164]	[0.252]	[0.116]
Observations	7,683	7,507	7,330	7,155	6,979	7,484	7,308	7,133	6,958	6,783
No. of country	182	182	182	182	182	179	179	179	179	179

Note: Dependent variable is log of GDP per capita while independent variable is log of total life expectancy. Columns 1-5 report dynamic fixed effects estimates while GMM estimates are denoted in columns 6-10 respectively. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses. All specifications include country and year fixed effects.
*p <0.10, **p <0.05, ***p <0.01.

In all specifications, columns 1-5, GDP is persistent and each estimate

$\Delta y_{it} = \theta_i y_{it-1} + \sum_l^q \gamma_{il} \Delta y_{it-l} + \epsilon_{it}$ (See Baltigi, 2005 p.254; Wooldridge 2002, ch.18).

is less than 1 indicating that estimates follow limiting distribution. We can consistently estimate β if GDP is stationary and health is uncorrelated with the factors in the error term. In dynamic within transformation, we removes time-invariant unobserved heterogeneity across countries by the time demeaning of original equation 1. The correlation of time-demeaned error terms decreases and eventually vanishes as T gets very large. Thus, we can achieve consistent estimates in dynamic within estimator as T tends to infinity.

Once we use the lagged values of the outcome variable as explanatory variables, the model suffers from Nickell's bias which varies with the order of $1/T$ where T is time period. This bias disappears if T tends to infinity.[9] In our case, each panel contains 56 time period; since our observations are very large, on average each panel has 42 observations, bias may be less. However, the error term contains time-varying factors which may affect simultaneously health and GDP per capita, they may lead our estimates biased and inconsistent. Thus, health is an endogenous variable. On the other hand, lagged GDP is also endogenous variable due to its correlation with the error term.

3.2 GMM Estimator

To address endogeneity of lagged outcome variables and health, we employ the first-differenced generalized method of moments (GMM) developed by Arellano and Bond (1991) in model 1 which also removes time-invariant factors.[10] Since

[9]Monte Carlo simulation studies explore that this bias declines as time period exceeds 20 (Papaioannou and Siourounis, 2008). Judson and Owen (1999) point out that bias is about from 1% to 2% of the true parameter when T is 30 and it is around 2% and 3% when T is 20.

[10]We use the dynamic IV estiamtor or Arellano-Bond GMM estimator or GMM estimator or differenced GMM interchangeably throughout the paper to refer to the same estiamtor.

y_{it-1} is correlated with the first-differenced error term, it cannot be considered as an instrument for the first differenced lagged outcome variables. However, y_{it-2}, y_{it-3},.., are uncorrelated with the error terms which are instruments for the first-differenced of lagged outcomes. On the other hand, health may be contemporaneously endogenous. So, h_{it-1} is no longer an instrument and h_{it-2}, h_{it-3},.., are instruments for the first-differenced of health. These internal instruments must satisfy the following orthogonal moments' condition in model 1 to identify β:

$$\mathbb{E}(\Delta u_{it}(h_{is}, y_{is})^{'}) = 0, \forall s \leq t-2$$

.

Columns 6-10, Table 2 reports the estimates from Arellano-Bond GMM estimation. The estimated effects of life expectancy are similar to dynamic within estimates (columns 1-5). The long-run effects are also comparable with within estimates. The preferred implied effects of life expectancy from GMM estimation is 3.954 (s.e=0.679) compared to 3.511 (s.e=0.917) in within estimate. The overall amounts of persistence in GDP are similar. Our GMM estimates are very similar to our preferred within estimates. To obtain consistent estimates, we need to test the assumption of no serial correlation in model 1: This assumption can be tested by the AR2 test. The p-value of this test is provided at the bottom which indicates that no autocorrelation assumption can be rejected for lags 2 (column 2) while we cannot reject for columns 1, 3, 4 and 5. So, our GMM estimates are consistent for lag 3.

Overall, our baseline results of dynamic within estimates of Table 2 show

a positive and significant relationship between income and health. In all specifications, columns 1 through 5, persistence in GDP is less than one which indicates that GDP is stationary. On the other hand, GMM estimates give us consistent estimates accounting for internal instruments for the lagged outcome and health.

The difference GMM produces too many instruments, the instrument count quadratic in T. This generates problem for finite samples and weaken the Hansen test. For this reason, we cannot report the number of instruments and Hansen p-value in Table 2 and onward. However, this does not affect our consistency of estimates. In this context, we need to reduce instruments and check the robustness of the results (Roodman, 2009). We reduces instruments substantially and find similar findings to our preferred estimates in robustness of subsection 3.3 (see Table A.1, Appendix A). Furthermore, in the same subsection, we account for various strategies to find the robustness of our results of Table 2 where both dynamic fixed effects and GMM estimators are considered.

3.3 Robustness

Model 1 captures the time-invariant factors by the country fixed effects while ignoring the time-varying factors which may be correlated with income and health simultaneously. We check robustness of our baseline estimates considering several strategies. We modify the model incorporating alternative health outcomes which are considered as determinants of income such as under-five child, adult male and maternal mortality rates on income. Table 3 reports

the effects of these measures. The estimates from all mortality rates are inversely related to income: An improvement in mortality leads to an increase in income. For example, under-five mortality is highly correlated with life expectancy; an improved under-five mortality increases some of life expectancy through increasing saving or reducing fertility and surviving children go to labor after long time and thus increase in income is much lower than preferred estimates: A 1% reduction in under-five mortality contributes in income by 0.408% (column 1, Table 3) compared to 3.511% (column 3, Table 2) in the long-run. However, maternal and adult mortality decline add more GDP per capita as they may enter labor market. For example, a 0.704% rises in income due to a 1% decreases in male adult mortality; however this is also less than our preferred estimates 3.511% (Column 3, Table 2) because adult mortality reduction increases some of life expectancy. The GMM estimates show that estimates are consistent with the estimates of baseline results.

Also, our model includes other health indicators such as male survival rate to age 65 which is a determinant of measuring GDP per capita (Well, 2007). Column 4 of Table 3 reports the effect of this factor on GDP per capita. Our results from this measure suggests that our baseline preferred findings remain similar.

Additionally, considering different subsamples based on different time regimes, we check robustness which is reported in Table 4. Column 1 of the table presents preferred estimates reproduced from baseline model taking into account all years to compare estimates. Columns 2-3 indicate the results from subsamples relying on year-intervals such as 1980-2015 and 1990-2015

respectively. The findings are close to preferred estimates. Panel B reports estimates from Arellano and Bond estimator and specifications in columns 2-3 produce similar to our preferred estimates (Column 1, Panel B).

Table 3: The effects of alternative health outcomes on GDP per capita

Health indicators	Child mort. under five (1)	Adult mort. male (2)	Mater. mort. (3)	Survival male (4)
		Panel A: Dynamic within estimates		
Health effects	-0.017***	-0.021***	-0.062***	0.045***
	(0.003)	(0.005)	(0.011)	(0.008)
GDP per capita	1.154***	1.073***	1.087***	1.035***
first lag	(0.062)	(0.053)	(0.087)	(0.052)
GDP per capita	-0.085	-0.031	-0.036	-0.030***
second lag	(0.069)	(0.024)	(0.084)	(0.023)
GDP per capita	-0.111***	-0.071**	-0.142***	-0.067***
third lag	(0.023)	(0.031)	(0.021)	(0.030)
Long run	-0.408***	-0.704***	-0.684***	1.750***
effect of health	(0.049)	(0.167)	(0.051)	(0.470)
Persistence	0.958***	0.971***	0.909***	0.974***
in GDP per capita	(0.005)	(0.014)	(0.005)	(0.005)
Observations	7,160	7,279	4,340	7,311
No. of country	183	181	176	176
		Panel B: GMM estimates		
Health effects	-0.023***	-0.032***	-0.094***	0.064***
	(0.004)	(0.010)	(0.019)	(0.012)
GDP per capita	1.127***	1.036***	1.023***	1.041***
first lag	(0.070)	(0.050)	(0.101)	(0.048)
GDP per capita	-0.081	-0.023	-0.028	-0.022
second lag	(0.068)	(0.022)	(0.080)	(0.022)
GDP per capita	-0.102***	-0.053**	-0.116***	-0.049**
third lag	(0.022)	(0.027)	(0.019)	(0.030)
Long run	-0.418***	-0.800***	-0.772***	2.196***
effect of health	(0.042)	(0.149)	(0.051)	(0.526)
Persistence	0.944***	0.960***	0.878***	0.971***
in GDP per capita	(0.009)	(0.008)	(0.024)	(0.008)
AR2 test (p-value)	[0.017]	[0.147]	[0.109]	[0.156]
Observations	6,977	7,094	4,164	7,135
No. of country	183	178	176	176

Note: The dependent variable is log of GDP per capita and independent variables are alternative health outcomes rather than life expectancy. Columns 1-3 denote under five child, male adult, and maternal mortality rates respectively while male survival to age 65 is indicated in column 4. All variables are in log form. We use our preferred lags in this Table: Three lags are used as a preferred lags in our analysis. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses. All specifications include country and year fixed effects.
*p <0.10, **p <0.05, ***p <0.01.

Furthermore, we check robustness of our results in model 1 by accounting for subsamples by regions (Table 5). The regions we consider are: Asia (South East

Asia, Asia Pacific and western Asia), Africa (Sub-Sahara and Arab), America (North America, Latin and Caribbean) and Europe (Europe and Central Asia). The estimates are reported in columns 2 through 5 of Table 5 respectively. Our findings are similar to preferred estimates even though estimates for all regions are larger except Africa region.

Table 4: The effect of life expectancy on GDP per capita with different time regimes

Year-intervals	60-15	80-15	90-15
	(1)	(2)	(3)
	Panel A: Dynamic within estimates		
Health effects	0.095***	0.182***	0.208***
	(0.016)	(0.027)	(0.032)
GDP per capita	1.070***	1.136***	1.132***
first lag	(0.052)	(0.080)	(0.087)
GDP per capita	-0.030	-0.054	-0.088
second lag	(0.023)	(0.083)	(0.100)
GDP per capita	-0.067**	-0.130***	-0.113***
third lag	(0.030)	(0.021)	(0.033)
Long-run effect	3.511***	3.731***	2.987***
of health	(0.917)	(0.398)	(0.295)
Persistence in	0.973***	0.951***	0.930***
GDP per capita	(0.005)	(0.007)	(0.009)
Observations	7,330	5,322	3,946
No. of country	182	182	182
	Panel B: GMM estimates		
Health effects	0.146***	0.308***	0.389***
	(0.024)	(0.076)	(0.103)
GDP per capita	1.034***	1.059***	1.024***
first lag	(0.048)	(0.093)	(0.120)
GDP per capita	-0.021	-0.040	-0.068
second lag	(0.022)	(0.079)	(0.099)
GDP per capita	-0.050**	-0.093***	-0.062***
third lag	(0.026)	(0.021)	(0.029)
Long-run effect	3.955***	4.172***	3.654***
of health	(0.679)	(0.423)	(0.406)
Persistence in	0.963***	0.926***	0.894***
GDP per capita	(0.007)	(0.018)	(0.027)
AR2 test (P-value)	[0.164]	[0.204]	[0.930]
Observations	7,133	5,128	3,759
No. of country	179	179	179

Note: Dependent variable is log of GDP per capita while independent variable is log of total life expectancy. Columns 1 presents estimates from 1960 to 2015 reproduced from Table 2 to make comparison. Columns 2-3 present estimates from 1980-2015, and 1990-2015 respectively. Standard errors robust to heteroskedasticity and autocorrelation at the country level are reported in parentheses. All models include country and year fixed effects.
*p <0.10, **p <0.05, ***p <0.01.

Table 5: The effects of life expectancy on GDP per capita with different regions

	Whole World (1)	Asia (2)	Africa (3)	America (4)	Europe (5)
		Panel A: Dynamic within estimates			
Health effects	0.095***	0.133**	0.079***	0.127***	0.271***
	(0.016)	(0.052)	(0.019)	(0.023)	(0.105)
GDP per capita	1.070***	0.995***	1.102***	1.003***	1.408***
first lag	(0.052)	(0.1592)	(0.086)	(0.024)	(0.066)
GDP per capita	-0.030	-0.005	-0.028	-0.002	-0.381***
second lag	(0.023)	(0.156)	(0.082)	(0.011)	(0.104)
GDP per capita	-0.067**	-0.010***	-0.111***	-0.020	-0.084
third lag	(0.030)	(0.024)	(0.033)	(0.013)	(0.052)
Long run	3.511***	6.784**	2.132***	6.795***	4.774***
effect of health	(0.917)	(2.179)	(0.732)	(2.596)	(1.269)
Persistence	0.973***	0.980***	0.963***	0.981***	0.943***
in GDP per capita	(0.005)	(0.009)	(0.009)	(0.006)	(0.009)
Observations	7,330	1,377	2,653	1,608	1,692
No. of country	182	35	63	35	49
		Panel B: GMM estimates			
Health effects	0.146***	0.133**	0.116***	0.141***	0.272 ***
	(0.024)	(0.052)	(0.024)	(0.026)	(0.103)
GDP per capita	1.034***	0.995***	1.072***	1.000***	1.408***
first lag	(0.048)	(0.159)	(0.087)	(0.023)	(0.064)
GDP per capita	-0.021	-0.005	-0.026	-0.002	-0.380***
second lag	(0.022)	(0.156)	(0.080)	(0.010)	(0.103)
GDP per capita	-0.050**	-0.010***	-0.094***	-0.019	-0.084*
third lag	(0.026)	(0.024)	(0.030)	(0.013)	(0.051)
Long run	3.955***	6.536***	3.004***	6.686***	4.774***
effect of health	(0.679)	(2.086)	(0.476)	(2.344)	(1.251)
Persistence	0.963***	0.980***	0.952***	0.979***	0.943***
in GDP per capita	(0.007)	(0.009)	(0.009)	(0.007)	(0.009)
AR2 test (p-value)	[0.164]	[0.178]	[0.064]	[0.011]	[0.378]
Observations	7,133	1,339	2,586	1,565	1,643
No. of country	179	34	63	33	49

Note: The dependent variable is log of GDP per capita and independent variable is log of life expectancy at birth. Column 1 covers the whole world which is produced from the Table 2 to compare with different subsamples. Columns 2-5 indicate estimates considering countries of Asia region (Southeast Asia, Asia-Pacific and Western Asia), Africa region (Sub-Sahara Africa and Arab), America region (North, Latin America and Caribbean countries), and Europe region (European and Central Asia) respectively. We use our preferred lags in this table: three lags are used as a preferred lags in our analysis. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses. All specifications include country and year fixed effects. *p <0.10, **p <0.05, ***p <0.01.

Moreover, we investigate the robustness of GMM estimates using alternative GMM estimator. While employing Arellano-Bond estimator we may have finite sample biased estimates due to instrumental proliferation. To overcome this problem, we employ alternative GMM estimator using truncated lags to 25. This reduces instruments largely, however our results remain similar to preferred

estimates of GMM (column 3, Appendix A, Table A.1).

Finally, our analysis deals with the presence of outliers reported in Appendix A, Table A.2 (columns 2-3). Excluding observations more and less than three standard deviations from mean, we estimate our preferred specification (column 2). Secondly, we account for outliers using Cook's distance where we exclude observations above a common rule-of-thumb threshold (four divided by the number of observations) (column 3). In both cases, results are very similar to our baseline findings which substantiate that our results are not influenced by outliers.

Overall, we use dynamic fixed effects estimators as well as Arellano-Bond GMM estimators: Both estimators provides very close results. Even though there have not been accounted for external instruments-which take into account the exogenous sources of variation in health-we consider robustness of our preferred estimates considering several time-varying controls both in dynamic within and GMM estimators. More importantly, panel data can sometimes capture the exogenous variation in institutions or policies through repeated observations of the same individual over long time period (Arellano, 2003). In addition, the panel data estimators we consider in this empirical research have asymptomatic properties: They can have consistent estimates as time period goes to infinity. Furthermore, the correlations of error terms-where time-varying factors contain-become less and less and eventually disappear as time period extends to infinity: This implies that estimates are consistent. Nonetheless, we intend to extend this discussion further to explore the asymptotic properties of dynamic estimators including external instruments in dynamic fixed effects

model in future.

4 Mechanisms and Test

As noted earlier, the existing literature has acknowledged the importance of mechanisms through which health affects income (e.g., Barro, 1996; Aghion et al., 2010; Zhang and Zhang, 2005). We investigate the channels and find that life expectancy has a significantly positive relationship with saving and negative relationship with fertility in the long-run.

To explore these channels or intermediate factors, we employ the following dynamic panel model:

$$c_{it} = \beta h_{it} + \sum_{l=1}^{q} \theta_l c_{it-l} + \alpha_i + \gamma_t + u_{it} \tag{2}$$

Where, c_{it} is one of possible channels that has effect on GDP. In this aspect, β indicates the impact of health on each channel. θ_l indicates coefficient of l^{th} lagged dependent variables of each channel which captures persistence in each channel. All other terms are the same as equation (1).

The channels we take into account include saving and fertility. Table 6 presents that life expectancy has a significantly positive effect on saving (column 1). However, it has a significantly negative effect on fertility (column 2).

Table 6: The effect of life expectancy on channels

	(1)	(2)	(3)
Health	1.562***	-0.033***	0.013
	(0.360)	(0.005)	(0.028)
Long run effect	10.391***	-1.572***	0.191
of health	(1.438)	(0.205)	(0.397)
Persistence	0.850***	0.979***	0.930***
in outcome	(0.019)	(0.002)	(0.007)
Saving			0.016***
			(0.002)
Fertility			-0.024***
			(0.006)
Observations	4,274	9,497	4,069
No. of country	169	182	137

Note: Dependent variables are channels while independent variable is log of life expectancy in columns 1-2 while column 3 indicates the test of mechanisms. Columns 1-2 report impacts of life expectancy on saving, and fertility respectively. Life expectancy, saving and fertility are in log form. Each channel is considered with 3 lags. In column 3, dependent variable is log of GDP per capita and independent variable is log of life expectancy and indicates estimates in the presence of channels such as saving, and fertility. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses. All specifications include country and year fixed effects.
*p <0.10, **p <0.05, ***p <0.01.

The presence of channels with life expectancy on the regression 1 leads to no effect of life expectancy on income suggests that our channels work on income. Column 3 of Table 6 demonstrates the test of mechanisms after considering all channels in equation 1. Results indicate that the relationship between life expectancy and GDP per capita vanishes when considering all channels together. Our channels such as saving is positive and significantly correlated with income while fertility is negatively correlated with income. Thus, this test proves empirically that our channels are potentials through which life

expectancy acts on income.

5 Conclusions

In this paper, we focus on the direct causal relationship between GDP per capita and life expectancy at birth using dynamic panel data models. Our paper employs the dynamic fixed effects estimator using a big longitudinal annual dataset over 1960-2015. This panel dataset covers the time regimes over 2000-2015 at which world leaders setup MDGs and health has substantially improved by reducing child mortality rates globally. Capturing the persistence in GDP per capita, we can consistently estimate parameter and models allow us to find the long-run effects of life expectancy at birth.

Our baseline results show that life expectancy has a significantly positive relationship with GDP per capita using the dynamic within estimator. When using lags as internal instruments in GMM estimator, the effects of life expectancy are similar to our preferred within estimates. Our analysis also explores mechanisms via which life expectancy acts on GDP per capita such as saving and fertility. The test of mechanisms provides evidence that life expectancy increases GDP per capita by decreasing fertility and increasing saving.

References

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

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):000–000.

Aghion, P., Howitt, P., and Murtin, F. (2010). The relationship between health and growth: when Lucas meets Nelson-Phelps. NBER working paper 15813, National Bureau of Economic Research, Cambridge, MA, USA.

Arellano, M. (2003). *Panel data econometrics*. Oxford university press.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Ashraf, Q. H., Lester, A., and Weil, D. N. (2008). When does improving health raise GDP? *NBER Macroeconomics Annual*, 23(1):157–204.

Baltigi, B. H. (2005). *Econometric Analysis of Panel Data*. John Wiley & Sons, Ltd.

Barro, R. J. (1996). Determinants of economic growth: a cross-country empirical study. NBER working paper 5698, National Bureau of Economic Research, Cambridge, MA, USA.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104:184–198.

Bhargava, A., Jamison, D. T., Lau, L. J., and Murray, C. J. (2001). Modeling the effects of health on economic growth. *Journal of Health Economics*, 20(3):423–440.

Chang, T., Chang, H.-L., Chu, H.-P., and Su, C.-W. (2006). Is per capita real GDP stationary in African countries? evidence from panel SURADF test. *Applied Economics Letters*, 13(15):1003–1008.

Cohen, D. and Leker, L. (2014). Health and education: Another look with the proper data. Discussion paper dp9940, Centre for Economic Policy Research, London, UK.

Cuestas, J. C. and Garratt, D. (2011). Is real GDP per capita a stationary process? smooth transitions, nonlinear trends and unit root testing. *Empirical Economics*, 41(3):555–563.

Ehrlich, I. and Lui, F. T. (1991). Intergenerational trade, longevity, and economic growth. *Journal of Political Economy*, 99(5):1029–1059.

Judson, R. and Owen, A. (1999). Estimating dynamic panel data models: a practical guide for macroeconomists. *Economic Letters*, 65(1):9–15.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). The worldwide governance indicators: methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2):220–246.

Meltzer, D. O. (1992). *Mortality decline, the demographic transition, and economic growth*. PhD thesis, University of Chicago, Department of Economics.

Papaioannou, E. and Siourounis, G. (2008). Democratisation and growth. *The Economic Journal*, 118(532):1520–1551.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population Studies*, 29(2):231–248.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *The stata journal*, 9(1):86–136.

Sachs, J. D. (2002). Macroeconomics and health: investing in health for economic development. *Revista Panamericana de Salud Pública*, 12:143–144.

Well, D. N. (2007). Accounting for the effect of health on economic growth. *The Quarterly Journal of Economics*, 122(3):1265–1306.

WHO and Others (2017). Health 2020. A European policy framework and strategy for the 21st century. *Cancer*.

Wooldridge, J. M. (2002). *Introductory econometrics: A modern approach*. The MIT Press.

World Bank (2018). World Development Indicators. Report, World Bank, Washington, D.C. Available from: http://databank.worldbank.org/ data//reports.aspx?source=2&country=&series =SH.DYN.MORT&period= [Accessed: 24 June 2018].

Zhang, J. and Zhang, J. (2005). The effect of life expectancy on fertility, saving, schooling and economic growth: theory and evidence. *The Scandinavian Journal of Economics*, 107(1):45–66.

Appendix A: Tables

Table A.1: The effect of life expectancy on GDP per capita with alternative GMM estimator

	(1)	(2)	(3)
Health outcome	0.095***	0.146***	0.202***
	(0.016)	(0.024)	(0.036)
Long-run effect	3.511***	3.955***	3.309***
of health	(0.917)	(0.679)	(0.577)
Persistence in	0.973***	0.963***	0.939***
GDP per capita	(0.005)	(0.007)	(0.012)
AR2 test (p-value)		[0.164]	[0.443]
No. of instruments		3,006	2,198
Observations	7,330	7,133	3,376
No. of country	182	179	147

Note: Dependent variable is log of GDP per capita while independent variable is log of total life expectancy. Column 1 presents preferred estimates in our baseline model. Columns 2 indicates preferred estimates from GMM estimator while column 3 denotes estimates using truncated lags to 25. All models include country fixed and year fixed effects. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses.
*p <0.10, **p <0.05, ***p <0.01.

Table A.2: The effect of life expectancy on GDP per capita

	(1)	(2)	(3)
Health	0.095***	0.090***	0.117***
	(0.016)	(0.015)	(0.017)
GDP per capita	1.070***	1.173***	1.153***
first lag	(0.052)	(0.063)	(0.068)
GDP per capita	-0.030	-0.099	-0.086
second lag	(0.023)	(0.073)	(0.076)
GDP per capita	-0.067*	-0.104***	-0.1002***
third lag	(0.030)	(0.024)	(0.023)
Long-run effect	3.511***	3.059***	3.559***
of health	(0.917)	(0.676)	(0.557)
Persistence	0.973***	0.971***	0.967***
in GDP	(0.005)	(0.004)	(0.005)
Observations	7,330	7,259	6,958
No. of country	182	178	178

Note: Dependent variable is log of GDP per capita while independent variable is log of total life expectancy. Column 1 presents preferred estimates in our baseline model. Columns 2 and 3 indicates estimates using our preferred specification. Column 2 denotes the estimates when excluding observations which have more or less standard deviation from mean. Column 3 shows Cook's distance when observations are above a common rule-of-thumb threshold. All models include country fixed and year fixed effects. Robust standard errors for heteroskedasticity at the country level and serial correlation are reported in parentheses.
*p <0.10, **p <0.05, ***p <0.01.

Appendix B: Figures

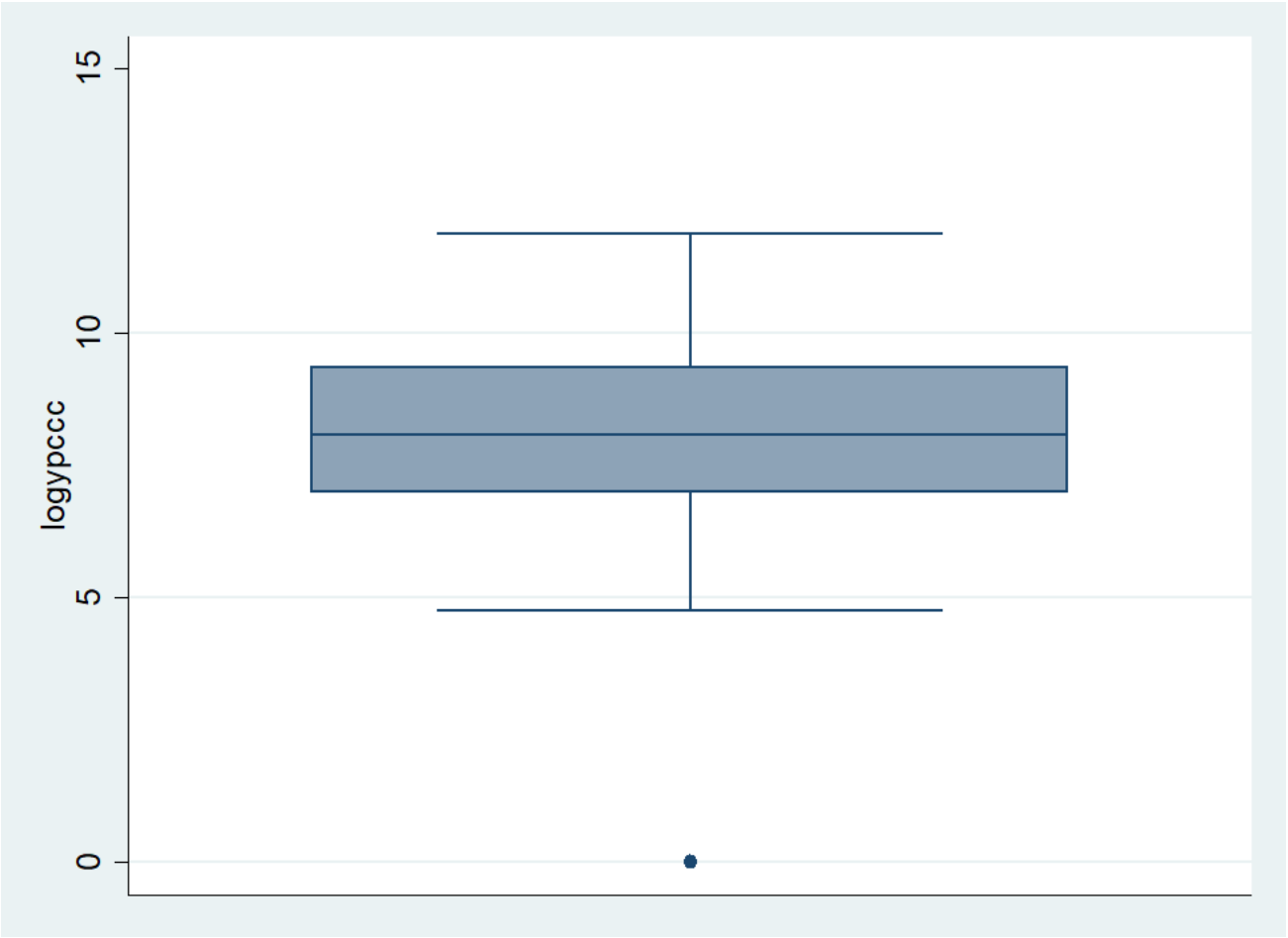

Figure.B1: The distribution of log (GDP per capita) with box diagram.

Note: logypccc denotes the log of GDP per capita (constant US dollar).

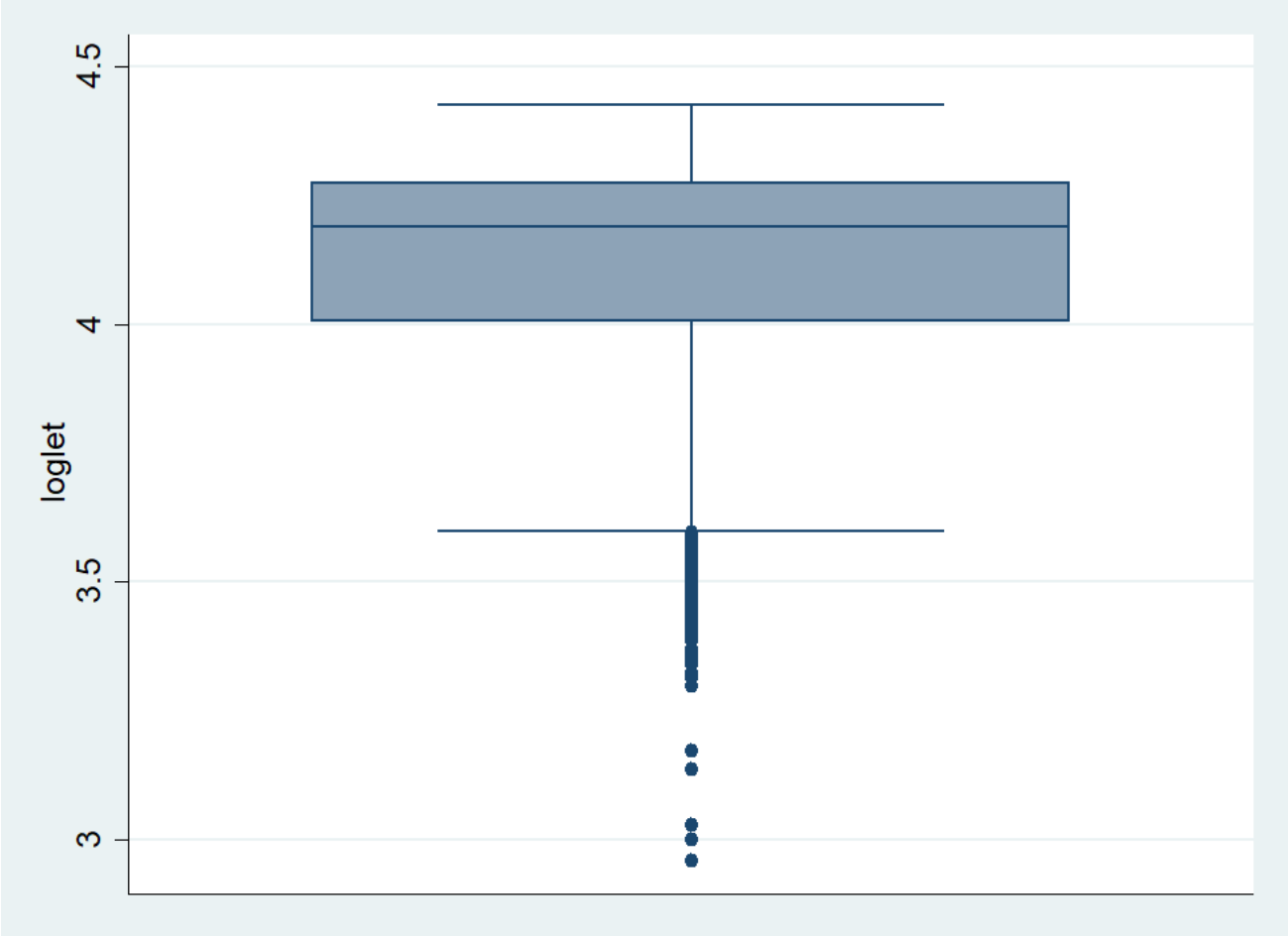

Figure.B2: The distribution of log (life expectancy at birth) with box diagram.

Note: loglet denotes the log of life expectancy at birth.