QUANTITATIVE METHODS FOR BUSINESS · GROUP PROJECT

Economic Development and Life Expectancy

Does higher income go together with longer lives? A multiple regression study of 179 countries using SPSS.

Multiple regression · Mediation · Moderation | IBM SPSS Statistics 29

RESEARCH QUESTION

Is higher GDP per capita associated with longer life expectancy across countries?

And does that relationship still hold once we control for other factors, such as public health spending, demographics and nutrition?

Dependent variable: life expectancy at birth (years), 2023
Main independent variable: GDP per capita, PPP-adjusted, 2023

HYPOTHESES

H_0 There is **no relationship** between GDP per capita and life expectancy.

H_1 Countries with **higher GDP per capita** have **higher life expectancy**.

How we tested it

Multiple regression, so the effect of GDP is measured **while holding the other variables constant**.

Significance level

$\alpha = 0.05$ (95% confidence). We reject H_0 if $p < 0.05$.

OUR DATA

179 countries, one year, 17 candidate variables

179
countries / territories in the dataset

17
candidate independent variables

150
complete cases in the final model

What we started with

- Economic: GDP per capita, trade, CO_2
- Health system: health spending, out-of-pocket share
- Living conditions: sanitation, electricity, internet, clean cooking
- Demographics & nutrition: fertility, undernourishment

Why N drops from 179 to 150

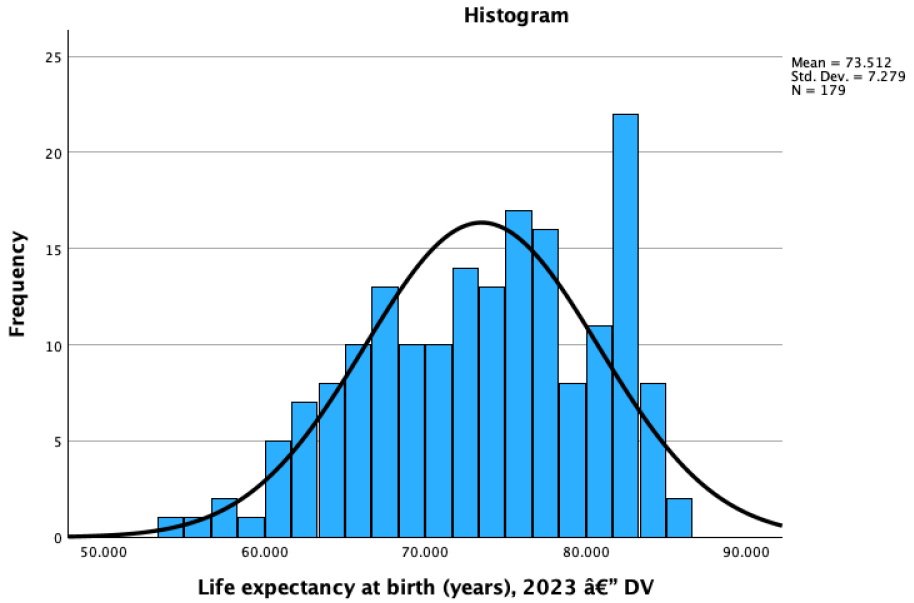
SPSS uses **listwise deletion**: a country is dropped if any variable in the model is missing. We did not impute values.

Source: public international indicator series (World Bank, UN, FAO), mostly 2023.

Is our dependent variable normally distributed?

Variable	Mean	Median	SD	Skew
Life expectancy (DV)	73.51	74.43	7.28	**-0.36**
GDP per capita	27,442	16,691	27,637	+1.49
Govt health spending	3.83	3.25	2.84	+1.70
Fertility	2.42	1.98	1.24	+1.08
Undernourishment	9.65	5.20	10.24	+1.98

Mean, median and skewness for the DV and the four variables we kept.

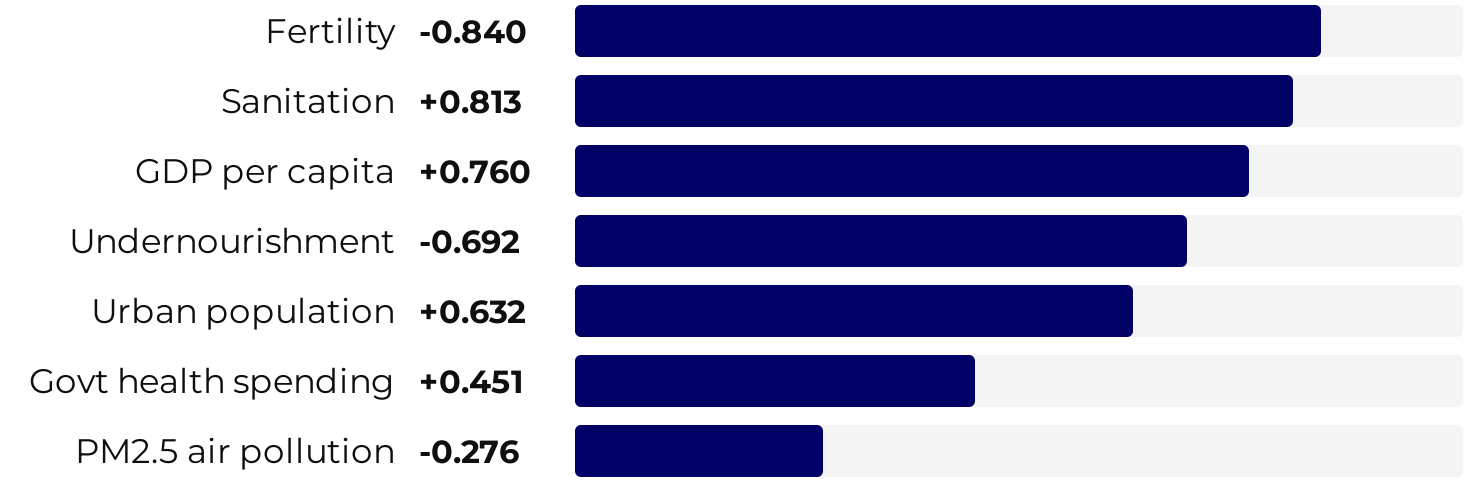


What we found

Roughly normal (skew **-0.36**), with a mild left tail and a cluster around 82–83 years. With N = 179 normality of the DV is not strictly required — but we checked, and it holds.

STEP 2 · CORRELATIONS

Which variables move with life expectancy?



Bar length = strength of the correlation, regardless of direction (+ or -).

		Life expectancy at birth (years), 2023 â€" DV
Life expectancy at birth (years), 2023 â€" DV	Pearson Correlation	--
	N	179
GDP per capita, PPP-adjusted constant international dollars, 2023	Pearson Correlation	.760**
	Sig. (2-tailed)	<.001
	N	179
Population using at least basic sanitation services (%), 2023	Pearson Correlation	.813**
	Sig. (2-tailed)	<.001
	N	179
Domestic general government health expenditure (% of GDP), 2023	Pearson Correlation	.451**
	Sig. (2-tailed)	<.001
	N	171
Total fertility rate (births per woman), 2023	Pearson Correlation	-.840**
	Sig. (2-tailed)	<.001
	N	179
Population living in urban areas (%), 2023	Pearson Correlation	.632**
	Sig. (2-tailed)	<.001
	N	179
Prevalence of undernourishment (%), 2023	Pearson Correlation	-.692**
	Sig. (2-tailed)	<.001
	N	153
PM2.5 mean annual exposure (micrograms per cubic meter), 2023	Pearson Correlation	-.276**
	Sig. (2-tailed)	<.001
	N	174

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

The life-expectancy column of our SPSS output.

All seven are significant (p < .001). So this test alone cannot tell us which four to use — it only tells us that nothing here is irrelevant.

From seven candidates down to four

7 candidates

Filter 1 — Is it related to life expectancy?

All seven pass. Everything correlates significantly with the DV, so nothing is dropped here.

5 left

Filter 2 — Does it tell us anything new?

Sanitation is cut: it correlates -.83 with fertility. **Urban population is cut:** +.64 with GDP. Both are near-copies of a variable we already have, so they add confusion, not information.

4 final

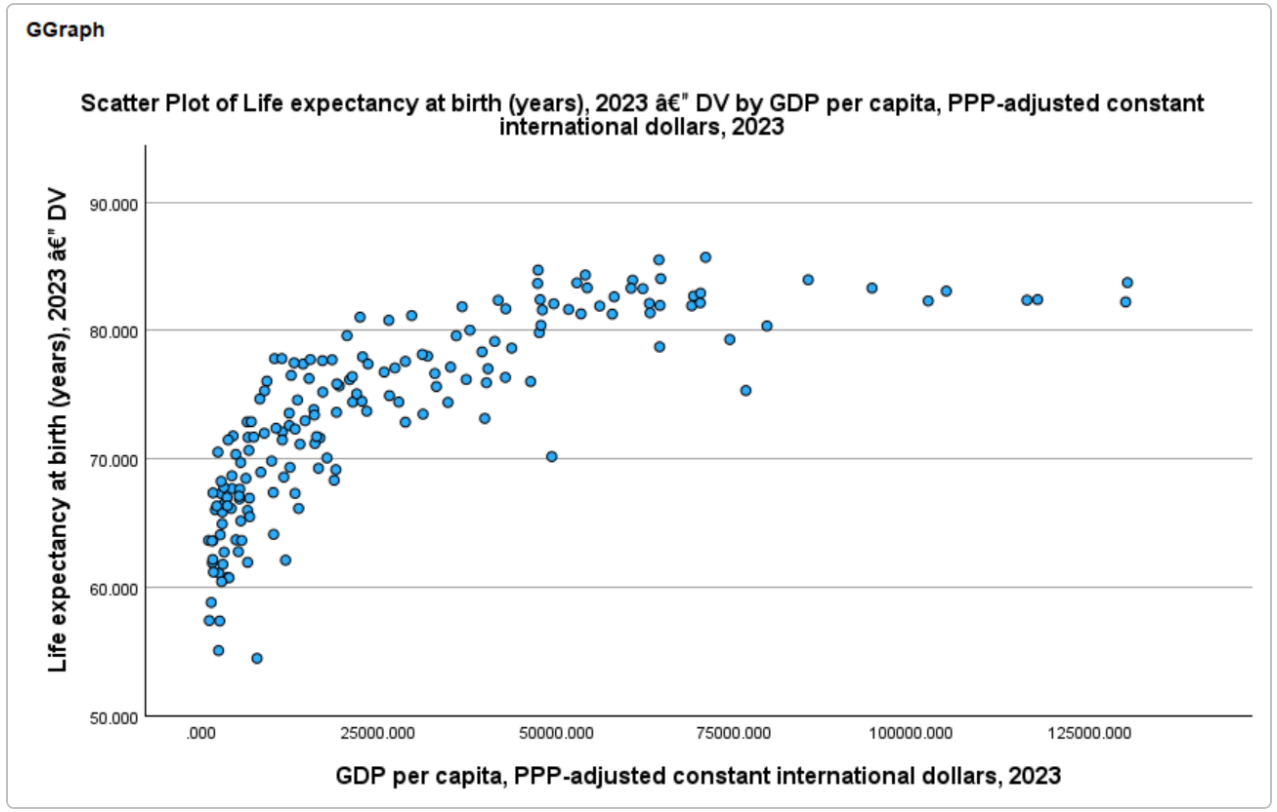
Filter 3 — Does it survive the regression?

PM2.5 is cut: significant on its own (-.28), but p = .627 once the other four are in the model. Its effect was already being captured by them.

The four we kept each represent a different channel: income (GDP per capita), demographics (fertility), nutrition (undernourishment) and the health system (government health spending).

STEP 3 · VISUALISING THE DATA

The GDP relationship is strong, but not a straight line



Life expectancy against GDP per capita, PPP-adjusted, 2023.

What we see

A clear upward relationship (r = .760), but the cloud **flattens out** above roughly $40,000 per capita.

What it means

Diminishing returns: an extra $1,000 buys far more life expectancy in a poor country than in a rich one. Income alone cannot be the whole story — which is exactly why we add controls in the next step.

STEP 4 · FIRST REGRESSION

We tested five variables and dropped one

Predictor	Standardised β	p	VIF	Decision
GDP per capita	+.318	<.001	2.16	**Keep**
Government health spending	+.118	.032	2.54	**Keep**
Fertility	-.492	<.001	2.60	**Keep**
Undernourishment	-.135	.007	2.07	**Keep**
PM2.5 air pollution	+.020	.627	1.50	Remove

Why PM2.5 was removed

p = .627 — it adds nothing once the others are in the model. R^2 stays at .83, and adjusted R^2 even improves slightly.

The principle

A simpler model with the same explanatory power is the better model. We removed it **with evidence**, not because we disliked the result.

Adding the variables one by one

Block	Model	R^2	Adjusted R^2	Change in R^2	p of the change
1	GDP only	.590	.587	+.590	<.001
2	+ Government health spending	.647	.642	+.057	<.001
3	+ Fertility	.820	.817	+.174	<.001
4	+ Undernourishment	**.830**	**.826**	+.010	.004

Every block adds something significant. GDP alone already explains 59% of the differences between countries; fertility adds the largest extra piece. This is why we can defend a four-variable model instead of simply entering everything at once.

STEP 6 · FINAL REGRESSION

Model fit and ANOVA

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.911[a]	.830	.826	2.965652	.830	177.407	4	145	<.001	1.975

a. Predictors: (Constant), Total fertility rate (births per woman), 2023, GDP per capita, PPP-adjusted constant international dollars, 2023, Prevalence of undernourishment (%), 2023, Domestic general government health expenditure (% of GDP), 2023

b. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6241.246	4	1560.311	177.407	<.001[b]
	Residual	1275.288	145	8.795		
	Total	7516.533	149			

a. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

b. Predictors: (Constant), Total fertility rate (births per woman), 2023, GDP per capita, PPP-adjusted constant international dollars, 2023, Prevalence of undernourishment (%), 2023, Domestic general government health expenditure (% of GDP), 2023

Result

R^2 = **.830** · F(4,145) = **177.4** · p **< .001**

We **reject the null hypothesis**: together the four predictors explain 83% of the differences in life expectancy between countries.

What each variable does

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	77.588	1.072		72.366	<.001	75.469	79.707		
	GDP per capita, PPP-adjusted constant international dollars, 2023	9.282E-5	.000	.318	6.336	<.001	.000	.000	.464	2.155
	Domestic general government health expenditure (% of GDP), 2023	.321	.149	.108	2.160	.032	.027	.615	.470	2.127
	Prevalence of undernourishment (%), 2023	-.096	.033	-.140	-2.915	.004	-.162	-.031	.507	1.974
	Total fertility rate (births per woman), 2023	-2.772	.309	-.487	-8.984	<.001	-3.382	-2.162	.398	2.511

a. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

In plain words

+$10,000 GDP per capita → about **+0.9 years**
+1 birth per woman → about **-2.8 years**
+1 pp government health spending → about **+0.3 years**
+10 pp undernourishment → about **-1.0 year**

Our hypothesis

GDP stays positive and significant even with all three controls in the model (**β = .318, p < .001**). All four predictors are significant at α = .05.

Is the model trustworthy?

Check	Result	Verdict
VIF (all predictors)	1.97 – 2.51	**No multicollinearity**
Tolerance	.398 – .507	**Acceptable**
Durbin-Watson	1.975	**Close to 2**
Std. error of estimate	2.97 years	**Typical error size**
Residuals	Bell-shaped, no pattern	**Assumptions OK**

VIF below 10 (here even below 3) means the predictors are not measuring the same thing.

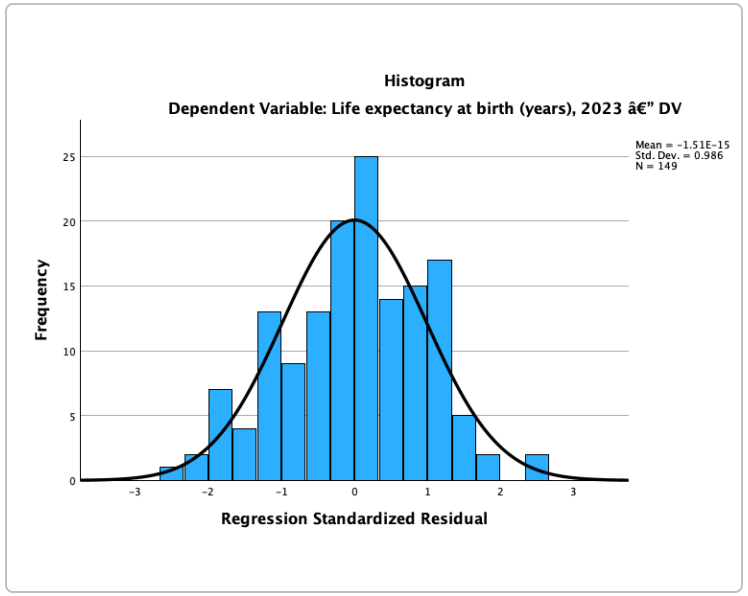
What VIF tells us

VIF above 10 would mean two predictors are measuring almost the same thing. Ours sit between 2 and 2.5, so each variable contributes its own information.

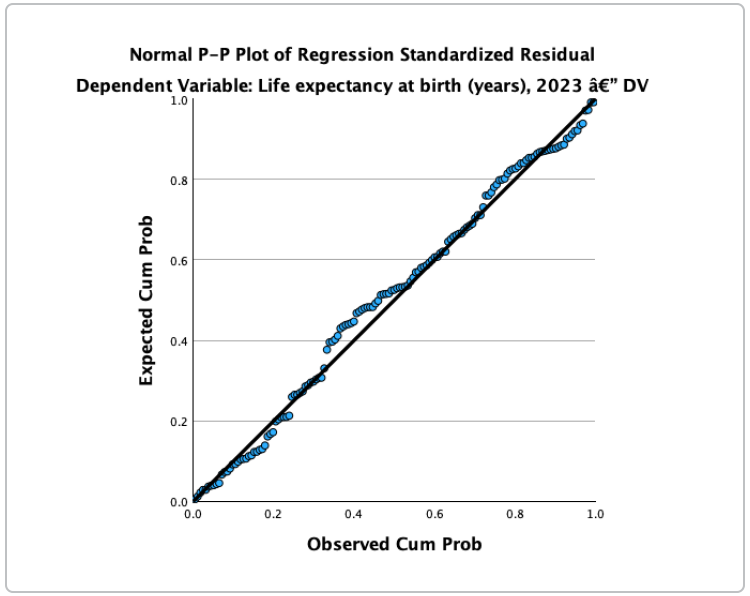
Outlier check

Nigeria is the largest outlier. Removing it barely changes anything (R^2 .830 → .836, same signs, still significant), so we **kept it**.

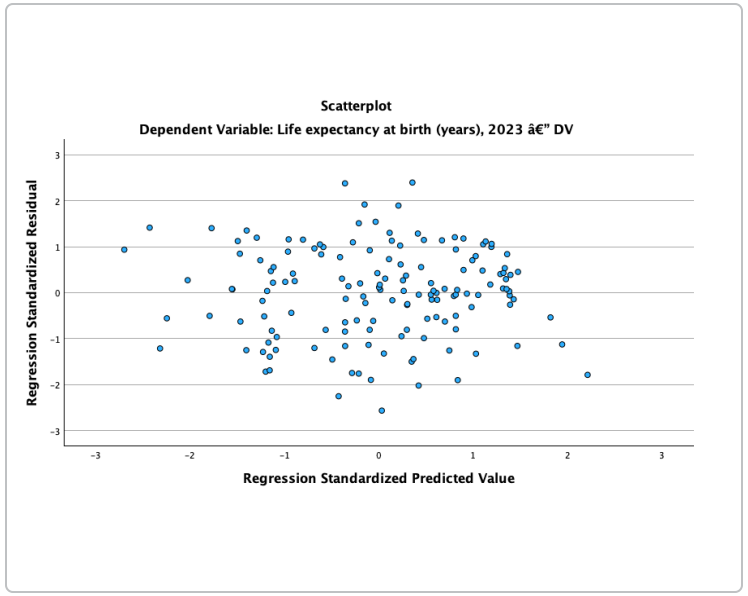
The residuals behave as they should



Histogram: bell-shaped around 0, SD = 0.986 → residuals are normal.



P-P plot: points follow the diagonal → normality confirmed.



Residuals vs predicted: random cloud, no funnel → homoscedasticity OK.

All three plots point the same way: the regression assumptions of **normal, evenly spread residuals** are satisfied, so the p-values and confidence intervals can be trusted.

Does GDP work partly *through* nutrition?

GDP per capita → **Undernourishment** → **Life expectancy**

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.770[a]	.593	.590	4.576342	.593	219.727	1	151	<.001	
2	.844[b]	.713	.709	3.853383	.120	62.975	1	150	<.001	2.076

a. Predictors: (Constant), GDP per capita, PPP-adjusted constant international dollars, 2023

b. Predictors: (Constant), GDP per capita, PPP-adjusted constant international dollars, 2023, Prevalence of undernourishment (%), 2023

c. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	67.737	.527		128.449	<.001	66.695	68.779
	GDP per capita, PPP-adjusted constant international dollars, 2023	.000	.000	.770	14.823	<.001	.000	.000
2	(Constant)	71.947	.692		103.999	<.001	70.580	73.314
	GDP per capita, PPP-adjusted constant international dollars, 2023	.000	.000	.564	11.082	<.001	.000	.000
	Prevalence of undernourishment (%), 2023	-.282	.036	-.404	-7.936	<.001	-.352	-.212

a. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

Partial mediation. GDP alone: β = .770. With undernourishment added it falls to β = .564 but stays significant — part of the GDP effect runs through nutrition, part does not.

STEP 10 · MODERATION

Does health spending change the GDP effect?

We centred both variables, built the interaction term **GDP × government health spending** and added it as a second block.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.911[a]	.830	.826	2.965652	.830	177.407	4	145	<.001	
2	.911[b]	.831	.825	2.972814	.000	.302	1	144	.583	1.976

a. Predictors: (Constant), Prevalence of undernourishment (%), 2023, centered value for Health, centered value for GDP, Total fertility rate (births per woman), 2023

b. Predictors: (Constant), Prevalence of undernourishment (%), 2023, centered value for Health, centered value for GDP, Total fertility rate (births per woman), 2023, gdp_x_health

c. Dependent Variable: Life expectancy at birth (years), 2023 â€" DV

Block 1 = four main effects. Block 2 adds the interaction: R^2 change = .000, Sig. F change = .583.

No moderation. The interaction is not significant (β = .023, p = .583); R^2 moves only from .830 to .831. We **fail to reject** this H_0 — a null result is still a result, not a failed analysis.

CONCLUSIONS

Is higher GDP per capita associated with longer life expectancy?

1

Yes, clearly

GDP stays positive and significant with all controls: $\beta = .318$, $p < .001$.

2

But not alone

Fertility is the strongest predictor ($|\beta| = .487$). Health spending and undernourishment add information beyond income.

3

Strong model

$R^2 = .830$, all predictors significant, VIF < 2.6, residuals well behaved.

4

Extensions

Nutrition partly mediates the GDP effect; health spending does not moderate it.

5

Limitations

One year of data, so no proof of causality; 29 countries lost to missing data; omitted variables remain possible.

Key takeaway: 83% explained variation is not 83% of the causes. Income is strongly associated with longevity, but demographics, nutrition and public health spending are part of the same story.