Panel Data Analysis of Health Expenditure and Life Expectancy at Birth of G20 Countries

Kowshik Sarker & Deeptam Chattopadhyay

03/05/2021

## Background

Human capital is considered an important factor for achieving desired economic growth and development in any country. Growth in human capital, in terms of education and health, positively affects per capita income in the long run. The current SARS CoV-2 outbreak has shaken the entire health system of the country and which compelled to rethink about the growth of health sector by expending more in it. In the due context, this study aims to analyze the relationship between dynamics of input and outputs of health care systems.

## Data and Variables

* We have collected the world development indicators data from World Bank Open Data [https://databank.worldbank.org/source/world-development-indicators#](https://databank.worldbank.org/source/world-development-indicators).
* We have considered the data for G20 countries from 2011 to 2020.
* Our exploratory variable is **Life Expectancy at Birt**h.
* Our explanatory variables are **GDP Growth**, **Current Health Expenditure**, **External Health Expenditure**

## Hypothesis

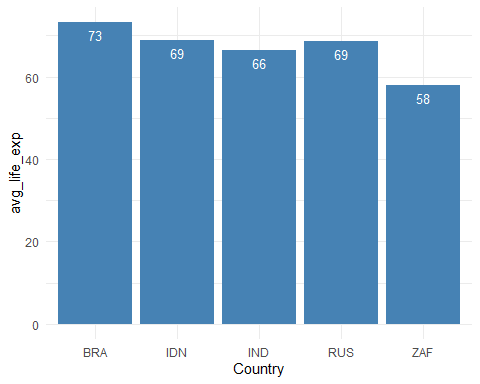
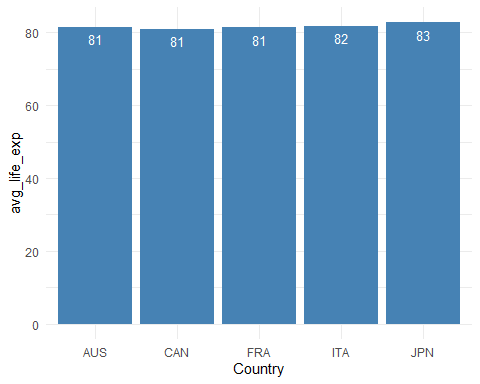
* Panel data modeling will give the best estimation for the dataset compared to POLS.
* Current health expenditure positively influence Life Expectancy.
* More the GDP Growth more will be the Current Health Expenditure
* External Health Expenditure has significant impact on Life Expectancy at Birth, thus providing strong point on FDI in health sector.

## Methodology

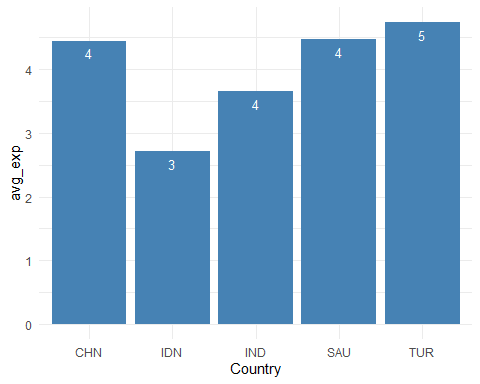
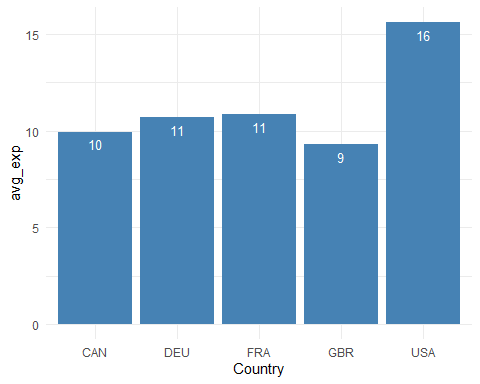
1. POLS (Pooled OLS)
2. Check for Heterogeneity both in Entity and Time.
3. Poolability Check
4. FEM (One Way & Two Way)
5. REM (One Way and Two Way)
6. Test for stability between FEM and REM
7. System GMM
8. Difference GMM
9. Panel Var
10. Final Model Comparison Based in Adjusted R-Squared and other Criterion.

## Basic EDA

## 'data.frame': 342 obs. of 6 variables:  
## $ Year : int 2001 2001 2001 2001 2001 2001 2001 2001 2001 2001 ...  
## $ Country : chr "IND" "AUS" "CAN" "SAU" ...  
## $ GDP\_Grwth : num 4.824 1.931 1.406 -1.211 0.998 ...  
## $ Curr\_Hlth\_Exp : num 4.26 7.7 8.66 4.46 13.22 ...  
## $ Ext\_Hlth\_Exp : num 0.461 0 0 0 0 ...  
## $ Life\_Exp\_At\_Brth: num 62.9 79.6 79.3 72.8 76.8 ...

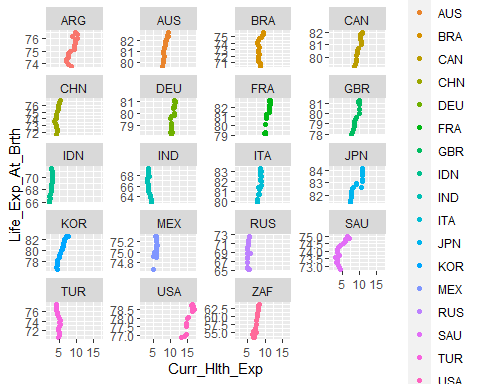


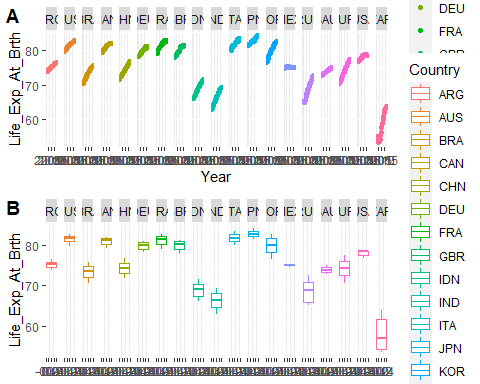
Avg Life Exp -Top 5 Country Avg Life Exp -Least 5 Country



Avg Health Exp -Top 5 Country Avg Health Exp -Least 5 Country

## Country wise Life Expectancy Plot trend

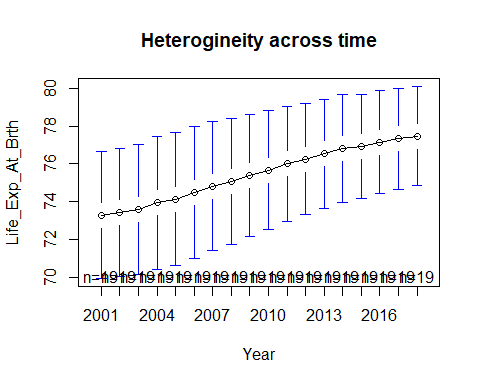
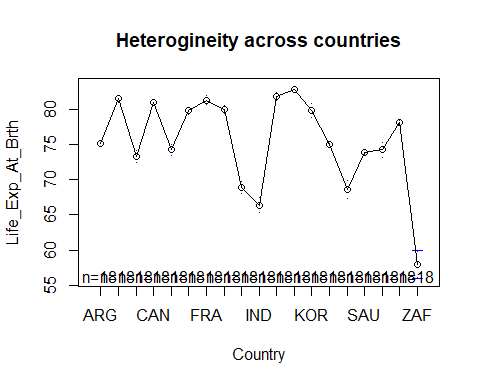




## POLS

## Pooling Model  
##   
## Call:  
## plm(formula = Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp +   
## Ext\_Hlth\_Exp, data = pdata, model = "pooling")  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -18.19728 -2.12697 0.75277 2.89648 24.25729   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## (Intercept) 69.946515 0.844343 82.8413 < 2e-16 \*\*\*  
## GDP\_Grwth -0.193174 0.080588 -2.3971 0.01707 \*   
## Curr\_Hlth\_Exp 0.958148 0.086621 11.0614 < 2e-16 \*\*\*  
## Ext\_Hlth\_Exp -1.280724 0.095725 -13.3791 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 14411  
## Residual Sum of Squares: 6887.2  
## R-Squared: 0.5221  
## Adj. R-Squared: 0.51786  
## F-statistic: 123.086 on 3 and 338 DF, p-value: < 2.22e-16

## Heterogeneity Check



Since the mean of the life expectancy is correlated with the countries and time, thus if we run POLS, there will be correlation between the error term and x, which present heteroskedasticity.

## One Way FEM

## Oneway (individual) effect Within Model  
##   
## Call:  
## plm(formula = Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp +   
## Ext\_Hlth\_Exp, data = pdata, model = "within")  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -4.251238 -0.705834 0.065101 0.640575 4.649138   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## GDP\_Grwth -0.040082 0.030368 -1.3199 0.1878   
## Curr\_Hlth\_Exp 1.010589 0.108936 9.2769 < 2.2e-16 \*\*\*  
## Ext\_Hlth\_Exp 0.252745 0.058802 4.2982 2.288e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 916.6  
## Residual Sum of Squares: 659.35  
## R-Squared: 0.28065  
## Adj. R-Squared: 0.23344  
## F-statistic: 41.6155 on 3 and 320 DF, p-value: < 2.22e-16

## Two Way FEM

## Twoways effects Within Model  
##   
## Call:  
## plm(formula = Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp +   
## Ext\_Hlth\_Exp, data = pdata, effect = "twoways", model = "within")  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -2.775336 -0.410930 0.014734 0.361979 3.701709   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## GDP\_Grwth -0.082758 0.023034 -3.5929 0.0003815 \*\*\*  
## Curr\_Hlth\_Exp -0.368521 0.089256 -4.1288 4.719e-05 \*\*\*  
## Ext\_Hlth\_Exp 0.152015 0.036900 4.1197 4.900e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 262.6  
## Residual Sum of Squares: 229.85  
## R-Squared: 0.12471  
## Adj. R-Squared: 0.014935  
## F-statistic: 14.39 on 3 and 303 DF, p-value: 8.6607e-09

## One Way REM

## Oneway (individual) effect Random Effect Model   
## (Swamy-Arora's transformation)  
##   
## Call:  
## plm(formula = Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp +   
## Ext\_Hlth\_Exp, data = pdata, model = "random")  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Effects:  
## var std.dev share  
## idiosyncratic 2.060 1.435 0.124  
## individual 14.526 3.811 0.876  
## theta: 0.9116  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -4.71064 -0.65861 0.21558 0.80251 3.91670   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)   
## (Intercept) 67.723651 1.233225 54.9159 < 2.2e-16 \*\*\*  
## GDP\_Grwth -0.040399 0.031384 -1.2873 0.1980026   
## Curr\_Hlth\_Exp 1.015789 0.105934 9.5889 < 2.2e-16 \*\*\*  
## Ext\_Hlth\_Exp 0.209093 0.060295 3.4678 0.0005247 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 1022.1  
## Residual Sum of Squares: 748.19  
## R-Squared: 0.26799  
## Adj. R-Squared: 0.2615  
## Chisq: 123.744 on 3 DF, p-value: < 2.22e-16

## Two Way REM

## Twoways effects Random Effect Model   
## (Swamy-Arora's transformation)  
##   
## Call:  
## plm(formula = Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp +   
## Ext\_Hlth\_Exp, data = pdata, effect = "twoways", model = "random")  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Effects:  
## var std.dev share  
## idiosyncratic 0.75858 0.87096 0.049  
## individual 14.59821 3.82076 0.947  
## time 0.05955 0.24402 0.004  
## theta: 0.9463 (id) 0.3665 (time) 0.3663 (total)  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -3.98614 -0.44007 0.22007 0.58730 3.55139   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)   
## (Intercept) 72.165761 1.474337 48.9479 < 2.2e-16 \*\*\*  
## GDP\_Grwth -0.065211 0.028989 -2.2495 0.0244805 \*   
## Curr\_Hlth\_Exp 0.440653 0.104208 4.2286 2.352e-05 \*\*\*  
## Ext\_Hlth\_Exp 0.193471 0.050539 3.8282 0.0001291 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 563.94  
## Residual Sum of Squares: 499.72  
## R-Squared: 0.11388  
## Adj. R-Squared: 0.10602  
## Chisq: 43.4393 on 3 DF, p-value: 1.9854e-09

## Poolability Test and Effect Test

##   
## F statistic  
##   
## data: Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp + Ext\_Hlth\_Exp  
## F = 10.65, df1 = 54, df2 = 266, p-value < 2.2e-16  
## alternative hypothesis: unstability

##   
## Lagrange Multiplier Test - two-ways effects (Gourieroux, Holly and  
## Monfort) for balanced panels  
##   
## data: Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp + Ext\_Hlth\_Exp  
## chibarsq = 1309.3, df0 = 0.00, df1 = 1.00, df2 = 2.00, w0 = 0.25, w1 =  
## 0.50, w2 = 0.25, p-value < 2.2e-16  
## alternative hypothesis: significant effects

* From the pooltest we can see that the same coefficients do not apply to each country. Thus the mean of each country is different across time.
* From plm test we can interpret that there is significant individual and time effects exists.

## Model Consistency Check

##   
## F test for twoways effects(BP Test)  
##   
## data: Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp + Ext\_Hlth\_Exp  
## F = 250.75, df1 = 35, df2 = 303, p-value < 2.2e-16  
## alternative hypothesis: significant effects

##   
## Hausman Test  
##   
## data: Life\_Exp\_At\_Brth ~ GDP\_Grwth + Curr\_Hlth\_Exp + Ext\_Hlth\_Exp  
## chisq = 251.52, df = 3, p-value < 2.2e-16  
## alternative hypothesis: one model is inconsistent

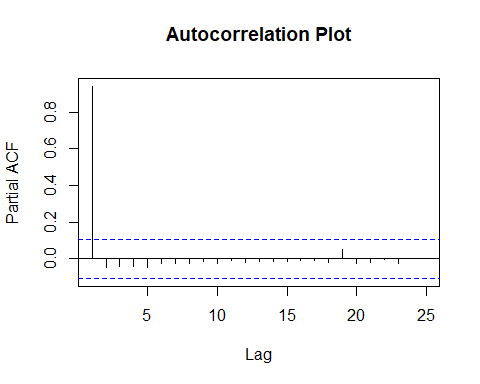
From both the test we can interpret that FEM and REM both are consistent enough to describe the individual and time effect better that POLS.

## Preliminary Model Summary

##   
## ====================================================  
## Dependent variable:   
## --------------------------------------  
## Life\_Exp\_At\_Brth   
## (1) (2) (3) (4)   
## ----------------------------------------------------  
## GDP\_Grwth -0.040 -0.083\*\*\* -0.040 -0.065\*\*   
## (0.030) (0.023) (0.031) (0.029)   
##   
## Curr\_Hlth\_Exp 1.011\*\*\* -0.369\*\*\* 1.016\*\*\* 0.441\*\*\*   
## (0.109) (0.089) (0.106) (0.104)   
##   
## Ext\_Hlth\_Exp 0.253\*\*\* 0.152\*\*\* 0.209\*\*\* 0.193\*\*\*   
## (0.059) (0.037) (0.060) (0.051)   
##   
## Constant 67.724\*\*\* 72.166\*\*\*  
## (1.233) (1.474)   
##   
## ----------------------------------------------------  
## Observations 342 342 342 342   
## R2 0.281 0.125 0.268 0.114   
## Adjusted R2 0.233 0.015 **0.261** 0.106   
## ====================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Causality test and Autocorrelation Check

##   
## Panel Granger (Non-)Causality Test (Dumitrescu/Hurlin (2012))  
##   
## data: Life\_Exp\_At\_Brth ~ Curr\_Hlth\_Exp  
## Ztilde = 11.704, p-value < 2.2e-16  
## alternative hypothesis: Granger causality for at least one individual



## System GMM

## System GMM

## Twoways effects Two steps model  
##   
## Call:  
## pgmm(formula = Life\_Exp\_At\_Brth ~ lag(GDP\_Grwth, 3) + lag(Curr\_Hlth\_Exp,   
## 3) + diff(Ext\_Hlth\_Exp, 3) | lag(Life\_Exp\_At\_Brth, 2:99),   
## data = pdata, effect = "twoways", model = "twosteps", collapse = TRUE,   
## transformation = "ld", index = c("Country", "Year"))  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Number of Observations Used: 551  
##   
## Residuals:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -21.489137 -0.625474 -0.002712 -0.499273 1.024596 7.768982   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)   
## lag(GDP\_Grwth, 3) -0.11129 0.15647 -0.7113 0.476919   
## lag(Curr\_Hlth\_Exp, 3) 1.03308 0.36257 2.8493 0.004381 \*\*  
## diff(Ext\_Hlth\_Exp, 3) -0.39278 0.95568 -0.4110 0.681072   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Sargan test: chisq(21) = 2.960051 (p-value = 1)  
## Autocorrelation test (1): normal = -0.4083162 (p-value = 0.68304)  
## Autocorrelation test (2): normal = 0.4580625 (p-value = 0.64691)  
## Wald test for coefficients: chisq(3) = 11.03523 (p-value = 0.011537)  
## Wald test for time dummies: chisq(14) = 64.33634 (p-value = 2.0088e-08)

## Difference GMM

## Twoways effects Two steps model  
##   
## Call:  
## pgmm(formula = Life\_Exp\_At\_Brth ~ diff(GDP\_Grwth, 1:1) + diff(Curr\_Hlth\_Exp,   
## 1:1) + diff(Ext\_Hlth\_Exp, 1:1) | lag(Life\_Exp\_At\_Brth, 2:99),   
## data = pdata, effect = "twoways", model = "twosteps", collapse = TRUE,   
## transformation = "d", fsm = "full", index = c("Country",   
## "Year"))  
##   
## Balanced Panel: n = 19, T = 18, N = 342  
##   
## Number of Observations Used: 304  
##   
## Residuals:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -0.723733 -0.136056 0.002240 0.008586 0.127261 0.959036   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)  
## diff(GDP\_Grwth, 1:1) 0.00062318 0.00535960 0.1163 0.9074  
## diff(Curr\_Hlth\_Exp, 1:1) -0.00284040 0.18380654 -0.0155 0.9877  
## diff(Ext\_Hlth\_Exp, 1:1) -0.00475318 0.03473883 -0.1368 0.8912  
##   
## Sargan test: chisq(16) = -4.848661e-21 (p-value = 1)  
## Autocorrelation test (1): normal = 1.228632 (p-value = 0.21921)  
## Autocorrelation test (2): normal = 1.975314 (p-value = 0.048233)  
## Wald test for coefficients: chisq(3) = 0.1379606 (p-value = 0.98692)  
## Wald test for time dummies: chisq(16) = 538.1701 (p-value = < 2.22e-16)

## Panel VAR

## 

## Final Model Comparison and Selection

