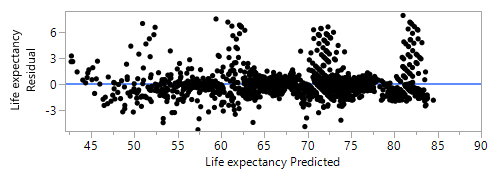
**panel data regression model analysis**

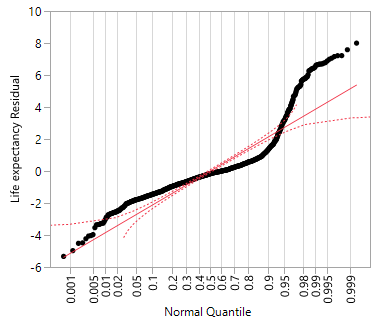
Assessing the use of panel data regression model

The first model to test was a multiple linear regression model. The assumptions for the use of this mode were tested and the results interpreted. The tests included test for normality, test for, independence, linearity, and equal variance assumption.

In the testing for normality assumption, the quantile plot of the residuals showed a an approximately linear plot. The normal quantile plot of residuals did not form linear diagonal line. This showed that the normality assumption was not satisfied by the data. While testing for the random assumption, the assumption could not be verified from the data because all the countries were sampled. For the independence assumption, the data collected are independent because they are collected from different countries for different measurements. Because of the independence sampling of data, independence assumption is satisfied. The assumption for equal variance was not satisfied because the plot of residuals against the predicted life expectancy is not spread uniformly around the line. Since some conditions were not met for using a multiple linear regression, the panel data regression model was the most preferred. Because of these violations, the multiple linear regression model was dropped in favor of the panel data regression model.



Plot of residuals against predicted.



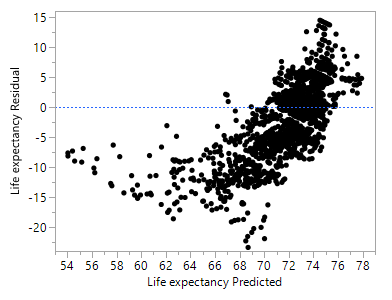
Normal quantile plot

The panel regression model is preferred because it takes into account both the random and fixed effect and minimizes bias. It can therefore create a better trend as compared to a multiple linear regression. Also, since the conditions for linear regression were violated, the panel regression that used an OSL model could not be used.

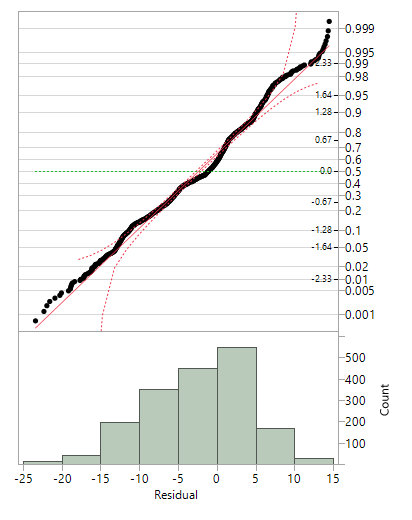
**Pane Data Regression Model**

The panel model appeared to be the best option for this study. The panel model is useful in creating trend in data that has been collected over time.

Testing assumptions for the model, the quantile plot of residuals formed a relatively linear diagonal line. This showed that the normality assumption was satisfied by the residual plot. This was also displayed by the histogram of the data. The histogram of the data appeared symmetrical. While testing for the random assumption, the assumption could not be verified from the data because all the countries were sampled. For the independence Assumption, the data collected are independent because they are collected from different countries for different measurements. Because of the independence sampling of data, independence assumption is satisfied. The assumption for equal variance was not satisfied because the plot of residuals against the predicted life expectancy is not spread uniformly around the line. Even though some conditions were not met, the model satisfied most conditions and can therefore be used.



A plot of Residuals against predicted



A normal quantile plot and a histogram of residuals.

A panel data has both a cross-section and a time-series dimension. In this data, all time-series observations are observed during the whole time period.

Xit,i=1,2,….n,t=1,.2…T T is generally small.

The standard static panel data model with i=1,…N ,t=1,…T is

Yit =β0 +x’itβ +€it.

Here, xit is a K-dimensional vector of explanatory variables without a constant term

β0 is the intercept which is not dependent on i and t.

β a (K×1) vector, the slopes are independent of I and t.

€it. , the error varies over I and t.

To successfully use this model technique, there are a few assumptions that are considered. The model can be designed to exhibits a constant coefficient for slope but has a changing intercept over time. The model can also be designed to have both the slope and intercept coefficients vary over time. Using these assumptions, models of fixed effect estimation and random effects estimation are created. These two models enable the panel regression model to take into account both time and trend factors in prediction. In this project, I used a fixed-effects model for the analysis because the fixed effects model is less biased and is not affected highly by the assumptions.

Reference

Cao, J. (2015, March 24). *Linear Mixed Models with JMP® Pro: One Face, Many Flavours*. Jmp.Com. https://community.jmp.com/t5/Discovery-Summit-Europe-2015/Linear-Mixed-Models-With-JMP-Pro-One-Face-Many-Flavours/ta-p/22333#:~:text=JMP%20Pro%2011%20has%20added,to%20its%20Fit%20Model%20platform.&text=Linear%20mixed%20models%20are%20a,%3DX%CE%B2%2BZ%CE%B3%2B%CE%B5 .

JMP. (n.d.). How do I build Repeated Measures and Panel Data Models? https://community.jmp.com/t5/Short-Videos/How-do-I-build-Repeated-Measures-and-Panel-Data-Models/ta-p/323564

JMP Help. (2021, May 5). Retrieved May 07, 2021, from https://www.jmp.com/support/help/en/16.0/index.shtml#page/jmp/mixed-models-and-random-effect-models.shtml#ww739671

Vella, F., & Verbeek, M. (1998). Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men. Journal of Applied Econometrics, 13(2), 163-183. Retrieved May 7, 2021, from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.18.3907&rep=rep1&type=pdf