**ECONOMICS SEMESTER-I PROJECT**

*Factors affecting GNI across countries*

Project Repo : [Sir practical\Sutanu\_ecotrix\_51.smcl](Sir%20practical/Sutanu_ecotrix_51.smcl)

SUTANU DAS

This project is part of the Semester course wherein we have tried to

determine the factors that can best explain the Gross National Income

seen across countries. In this pursuit, we have incorporated the use of OLS

regression theory to determine a suitable model.

The procedures followed have been chosen to look for discrepancies in my

data that may affect the model conclusions.

The data used in this project consists of observations related to different

aspects of a country’s development collected across 217 countries, for the

year 2015.

**METHODOLOGY**

My main objective was to see if the factors such GDP, GNI per capita, inflation, foreign exchange and the other explanatory variable could explain the variations of Gross National Income across countries. In the given data I have taken all possible variables that both direct and indirect indicators which could have effect on Gross National Income.

Dependent Variable – Gross National Income (GNI)

Independent Variables – Gross National Income per capita

Current Gross Domestic Product

Annual GDP growth

GDP per capita in USD

Inflation (GDP deflator) 2015

Inflation (GDP deflator) 2010

Gross Capital Formation

Debt service as a percentage of GNI

Net foreign direct investment inflows

Net foreign direct investment

Inflation (consumer prices) 2015

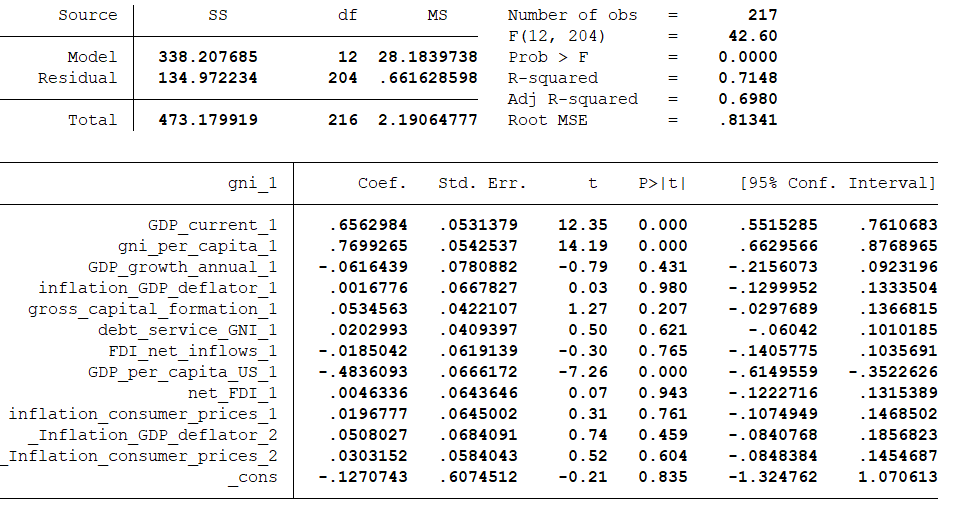
Inflation (consumer prices)2010

All of these variables are directly associated in explaining Gross National Income.

**ANALYSIS DONE USING STATA**

The OLS regression was done in STATA using the following command :-

**regress gni\_1 GDP\_current\_1 gni\_per\_capita\_1 GDP\_growth\_annual\_1 inflation\_GDP\_deflator\_1 gross\_capital\_formation\_1 debt\_service\_GNI\_1 FDI\_net\_inflows\_1 GDP\_per\_capita\_US\_1 net\_FDI\_1 inflation\_consumer\_prices\_1 \_Inflation\_GDP\_deflator\_2 \_Inflation\_consumer\_prices\_2**



The summary statistics suggests that GDP current, GNI per capita, inflation GDP deflator 2015, Gross capital formation, debt service GNI, net FDI, inflation consumer prices 2015, inflation GDP deflator 2010 all these have positive coefficients which suggests that all of these explanatory variables have positive impact on the level of GNI across countries.

GDP growth annual, FDI net inflows, GDP per capita US, inflation consumer prices 2010 have negative coefficients suggesting that all these explanatory variables have negative impact on the dependent variable.

The overall fit of the model is quite moderately good as indicated by high values of R2

and adjusted R2, which are 0.71 and 0.69 respectively.

**Model Diagnostics: Checking for violations of Assumptions**

The methodology was followed based on the assumptions of the Classical

OLS model. We now check for the violations of some of these assumptions.

**1. Violation of Homoscedasticity- Presence of Heteroscedasticity**

We first proceed with visual inspection of the data to determine the presence

of heteroscedasticity by plotting the residuals against the fitted values. We

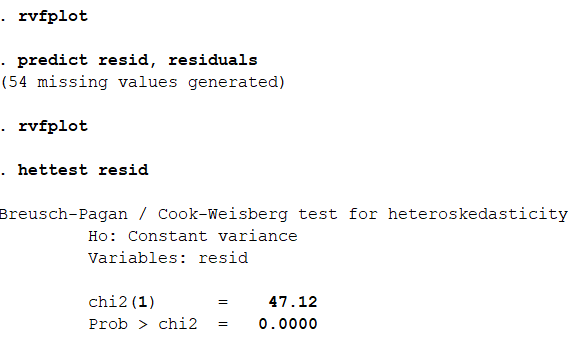
see there may be some presence of heteroscedasticity. We try confirm our

suspicions by running the Breusch-Pagan / Cook-Weisburg test for residuals

which returns a p-value less than 0.05.

Thus, we may conclude our data does not have presence of

heteroscedasticity.





**2. Violation of normality assumptions**

From visual inspection of the quantile-quantile plot of the fitted residuals

against the normal distribution we find there may be a possibility that the

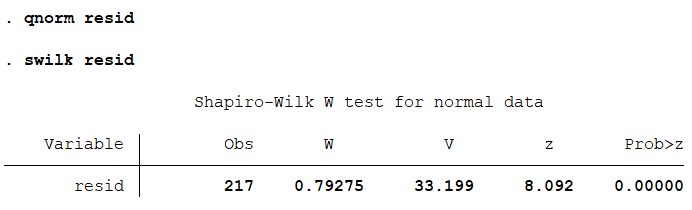
normality assumption is violated. The fitted residuals do not perfectly coincide

with the normality line.

However, on conducting the Shapiro-Wilks test the p-value is quite low

and is below our 5% probability of rejection. So, we conclude that the

distributional assumption of normality still holds.



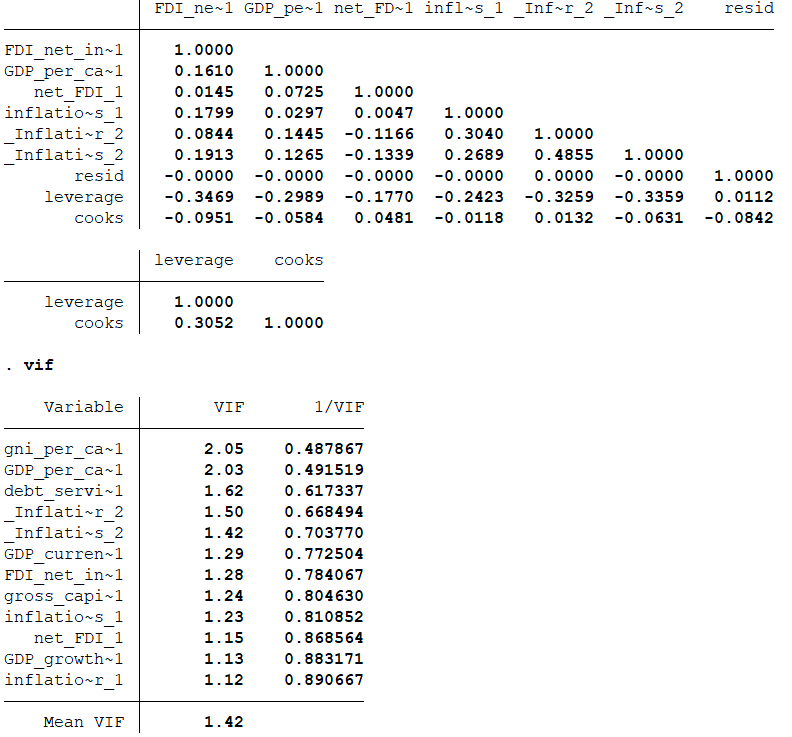


**3. Violation of multicollinearity**

The pairwise-correlation matrix helps with the visual verification of the presence of

multicollinearity. We find there to be some minor correlation between the regressors.

The Variance Inflating Factor is found to see how much this correlation contributes to

the OLS coefficients. It is found all of them are quite close to 1, indicating that the correlation present does not pose a major obstacle in our analysis. 

**4. Testing for linearity of model.**

We try to confirm visually if the dependent and independent variables

have a linear relationship via individual scatter plots.













**5.Alternative measures of goodness of fit.**

Aside from the standard measures of goodness of fit R2 and adjusted R2 we

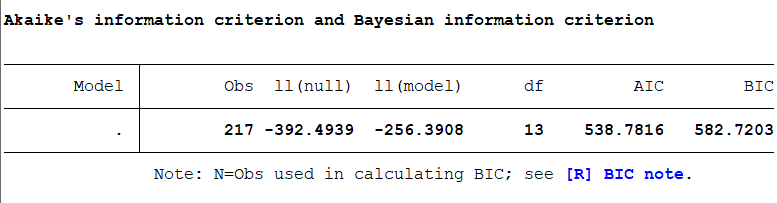
try and see if our conclusions about the original model have and the new

model changes when we use other measures.

1.) Akaike Information Criteria

2.) Bayesian Information Criteria

Both criteria impose a penalty on adding additional regressors and so do not increase with the addition of new variables like the standard R2. Further the penalties imposed mean these measures put a premium on achieving a better fit with a smaller number of regressors.



In both cases, we find AIC and BIC have quite high values, indicating our models are good fits and can explain the variations in the data.

**6. Influence Statistics and their interpretation**

A scatter diagram of leverage, residuals, and Cook's distance can help identify influential data points in a regression model:

* Cook's distance

A measure of how much the model's estimates change when a specific observation is removed. A Cook's distance value above 1.0 indicates a highly influential point. Values above 0.5 indicate some influence, and values below 0.5 indicate no influence.

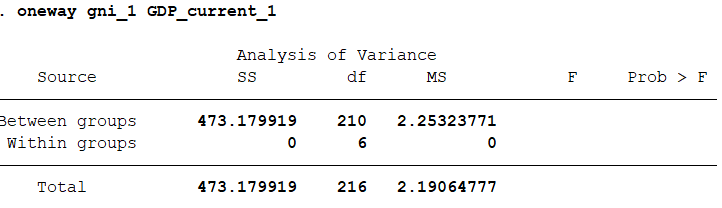
* Leverage

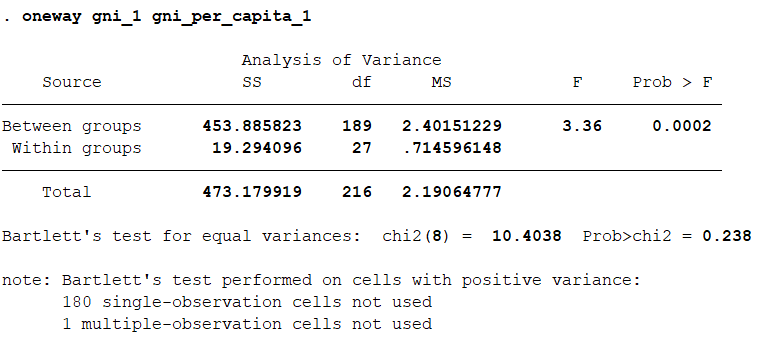
A measure of how far a predictor value is from its average. High leverage observations have predictor values that are far from their averages which greatly influence the fitted model.

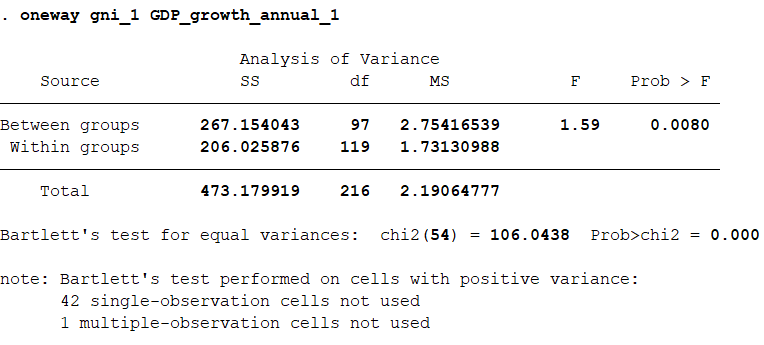


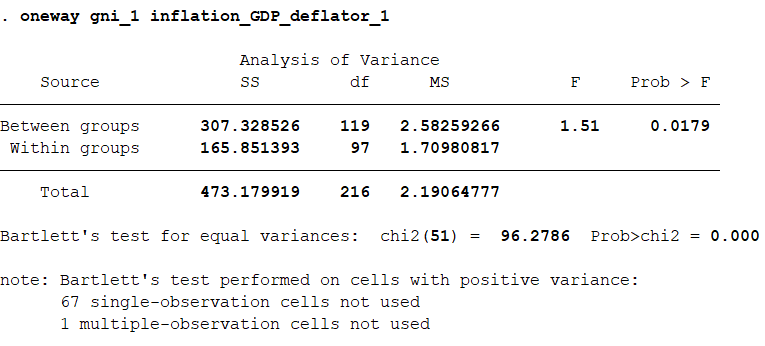
**7. Tests of Hypotheses: Theoretical Hypotheses and Testing for Relation among parameters. Model Comparison under Null with Unrestricted Model using ANOVA**

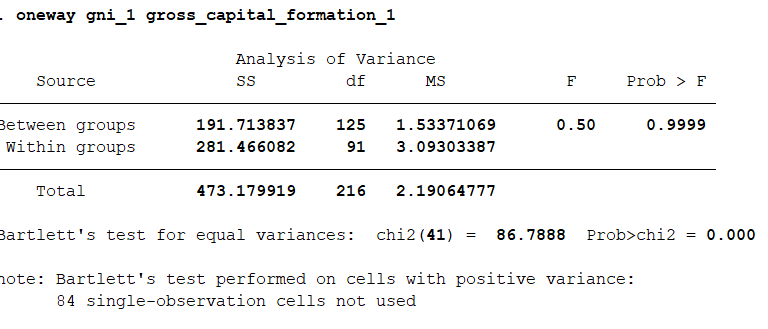
The ANOVA test suggests that in the model the additional predictors do not improve model performance.

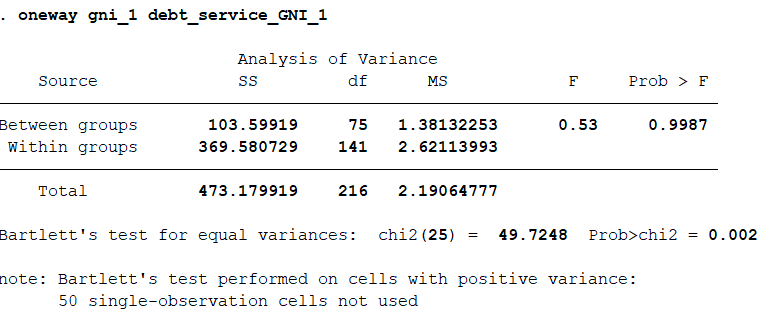
****

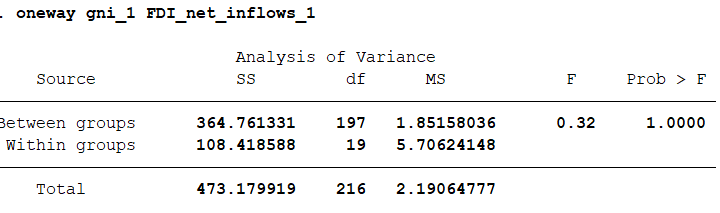
****

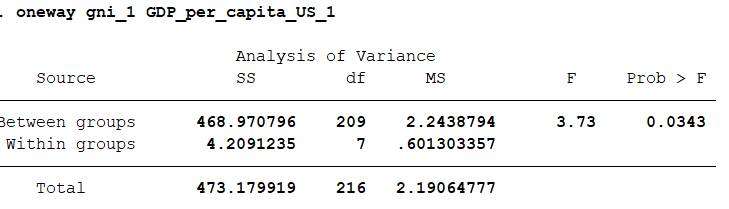
****

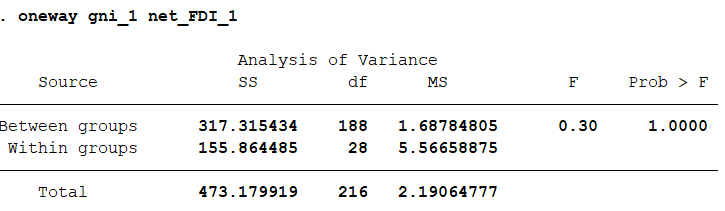
****

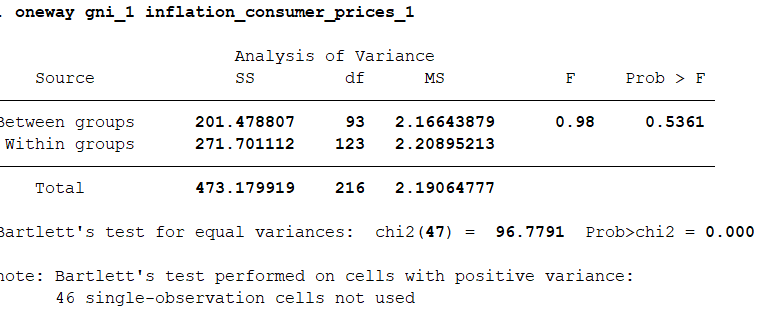
****

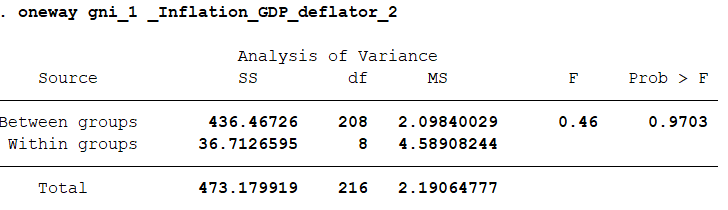
****

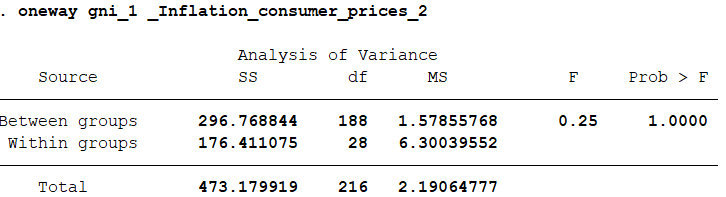
****

****

****

****

****

****

**CONCLUSION:-**

The model we have chosen is quite adequate in explaining the variations in Gross National Income seen across countries. We can conclude that the variables Gross National Income per capita, Current Gross Domestic Product, Annual GDP growth, GDP per capita in USD,

Inflation (GDP deflator) 2015, Inflation (GDP deflator) 2010, Gross Capital Formation,

Debt service as a percentage of GNI, Net foreign direct investment inflows,

Net foreign direct investment, Inflation (consumer prices) 2015,

Inflation (consumer prices)2010 have significant roles in determining the Gross National Income of the country.

Project Repo : [Sir practical\Sutanu\_ecotrix\_51.smcl](Sir%20practical/Sutanu_ecotrix_51.smcl)