QMBU 450 Final Project

Relationship Between Good Governance and Inequality

Bülent Söyler | ID: 54303

İsmail Berk Sabuncuoğlu | ID: 53623

INTRODUCTION

In this paper, the results of a data analysis are summarized, which is conducted to explore a phenomenon that is related to the field of Economic Development: The relationship between the quality of governance and inequality. The authors of the paper believe that the good governance of a country and inequality are negatively correlated. For that, the significance levels between the two is examined. Hypothesis created based on the subject is:

Hypothesis: There is a significant relationship between the good governance and inequality of a country.

In this paper, the accuracy of the hypothesis is examined.

VARIABLES

For the analysis, explanatory and dependent variables were needed. Two indicators related to good governance and inequality were chosen: World Governance Indicators (WGI) and GINI Index

World Governance Indicators (WGI): WGI is calculated by the World Bank since 1996 to measure how good the countries are governed. The calculations are based on the perceptions gathered from multiple civil rights organizations and public institutions. WGI was conducted once in two years until 2002 for over 200 countries. After 2002, it is conducted annually. According to the World Bank, WGI captures six key dimensions of governance:

- **Voice and Accountability:** Reflects the perceptions of the extent of participation of a country's citizens to choose their government as well as other freedoms like freedom of speech, freedom of press, and so forth.
- **Political Stability and Absence of Violence/Terrorism (Political Stability No Violence):** Reflects the perceptions of likelihood of political stability as well as violence that is politically motivated, including terrorism.
- **Government Effectiveness:** Reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.
- **Regulatory Quality:** Reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.
- **Rule of Law:** Reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.
- **Control of Corruption:** Reflects perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests (*World Bank, 2019*).

Measurement: The values for each dimension range from -2.5 (weak governance) to 2.5 (strong governance)

GINI Index: GINI Coefficient, or GINI Index, is a statistical measurement represents the income or wealth distribution of a nation's residents. It is developed by the Italian statistician and sociologist Corrado Gini in 1912 and it is the most widely used measurement of inequality (*Chappelow, 2020*).

Measurement: The GINI coefficient for a country range from 0 (perfect equality) to 100 (perfect inequality).

DATA

For gathering the data, two Excel files were downloaded from the official website of the World Bank. One of them is the data of WGI from the year 1996 to 2018 and the other one is the data of GINI Index from the year 1960 to 2019. Both files have the data for over 200 countries. The files can be found from these websites:

- For WGI: https://info.worldbank.org/governance/wgi/
- For GINI Index: https://data.worldbank.org/indicator/SI.POV.GINI

MODELS AND FINDINGS

For indicating whether the hypothesis can be rejected or not, OLS Regression analysis with seven variables was conducted. The six dimensions of WGI were the explanatory variables and GINI coefficient was the dependent variable. The variables were named in the model as like this:

- **GINI:** GINI Index
- **GE:** Government Effectiveness
- **VA:** Voice and Accountability
- **CoC:** Control of Corruption
- **PSNV:** Political Stability No Violence
- **RQ:** Regulatory Quality
- **RoL:** Rule of Law

Since there is a high chance the explanatory variables are correlated, Principal Component Analysis (PCA) was also used in the analysis.

Although the WGI values for every single country were available, the GINI data did not contain the values for every single country. For that, the countries who did not contain GINI coefficient for any year were eliminated. When the elimination process was finished, 160 countries were left for observation, and the two files on Excel were merged and pulled as a one file via Python (Spyder). The first ten observations of the data are shown below:

Countries	GINI	VA	CoC	GE	PSNV	RQ	RoL
Albania	31,78571	0,100883	-0,635056948	-0,2724	-0,05382	0,061584	-0,53057
Algeria	27,6	-0,92926	-0,599621999	-0,53831	-1,20368	-0,8879	-0,75258
Angola	47	-1,15498	-1,317454506	-1,12747	-0,5664	-1,02999	-1,28438
Argentina	45,0125	0,373669	-0,390413581	-0,10186	-0,07998	-0,67456	-0,61653
Armenia	31,96471	-0,60941	-0,609088721	-0,13203	-0,12215	0,227336	-0,36942
Australia	34,22	1,414788	1,929785118	1,702943	0,956616	1,657986	1,777569
Austria	30,26667	1,386632	1,709091032	1,69054	1,161732	1,435733	1,863474
Azerbaijan	26,325	-1,31004	-1,056557613	-0,58396	-0,71158	-0,39635	-0,77972
Bangladesh	32,56667	-0,48757	-1,085489939	-0,76095	-1,35091	-0,86742	-0,82655

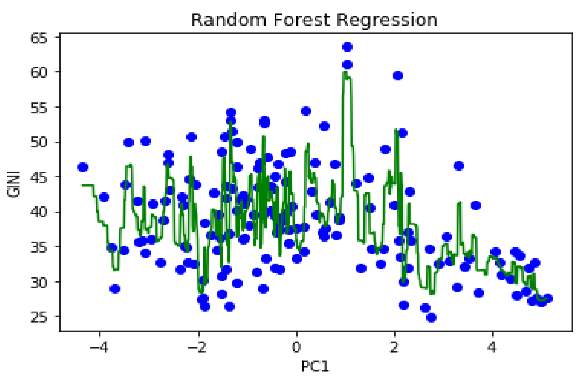
The values of the variables for each country were calculated via taking the average of the values in the 2002-2018 period for each variable. Three different models were created for the analysis.

Model 1: With 160 observations, the model was conducted via PCA and OLS Regression. For PCA, the explanatory variables were standardized, the covariance matrix was computed, and it was used to compute the Eigenvalues and Eigenvectors. With PCA, a variable named “PC1” was created with the six WGI variables.

Linear Regression was used for both PC1 and the six WGI variables for comparison. Random Forest Regression (RFR) was also used for PC1 and GINI, since RFR is not linear and can give more accurate results. First, an RFR graph was created with PC1 and GINI. After that, the data was divided into training and test data and the Mean Absolute Error (MAE) Mean Square Error (MSE), Root Mean Square Error (RMSE) and R2 values were calculated for both training and test. The training data consists of 112 observations and the test data consists of 48 observations. The results are shown below:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.080
Model:                            OLS   Adj. R-squared:                  0.074
Method:                 Least Squares   F-statistic:                     13.78
Date:                Tue, 19 May 2020   Prob (F-statistic):           0.000284
Time:                        01:48:45   Log-Likelihood:                -546.83
No. Observations:                 160   AIC:                             1098.
Df Residuals:                     158   BIC:                             1104.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     38.8404      0.587     66.157      0.000      37.681      40.000
PC1           -0.9534      0.257     -3.713      0.000      -1.461      -0.446
==============================================================================
Omnibus:                       14.950   Durbin-Watson:                   1.830
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               16.486
Skew:                           0.691   Prob(JB):                     0.000263
Kurtosis:                       3.749   Cond. No.                         2.29
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.207
Model:                            OLS   Adj. R-squared:                  0.176
Method:                 Least Squares   F-statistic:                     6.653
Date:                Tue, 19 May 2020   Prob (F-statistic):           2.85e-06
Time:                        01:48:45   Log-Likelihood:                -534.97
No. Observations:                 160   AIC:                             1084.
Df Residuals:                     153   BIC:                             1105.
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     38.4731      0.570     67.471      0.000      37.347      39.600
VA             3.4027      1.318      2.582      0.011       0.799       6.006
CoC            4.9697      2.253      2.206      0.029       0.519       9.421
GE            -2.0554      2.691     -0.764      0.446      -7.372       3.261
PSNV           1.4162      1.059      1.337      0.183      -0.677       3.509
RQ             1.3300      2.294      0.580      0.563      -3.203       5.863
RoL          -10.5578      2.908     -3.630      0.000     -16.304      -4.812
==============================================================================
Omnibus:                        9.351   Durbin-Watson:                   1.847
Prob(Omnibus):                  0.009   Jarque-Bera (JB):               10.679
Skew:                           0.431   Prob(JB):                      0.00480
Kurtosis:                       3.926   Cond. No.                         13.1
==============================================================================
```



With RFR:

Mean Absolute Error (MAE): 6.523316557519109

Mean Squared Error (MSE): 67.55885880128454

Root Mean Squared Error (RMSE): 8.219419614625144

R2 (Training data): 0.8581367676369575

R2 (Test data): -0.2551940732963556

Although that the probability of F-statistic is less than 5%, both SLR with PC1 and MLR with the six WGI variables cannot explain the most of variability in GINI, since their Adjusted R2 levels are very low. Also, the P-values of GE, PSNV, and RQ are higher than 5%, which shows that there is a multicollinearity between the values, and this can be shown via correlation table that shows high correlation between the values below:

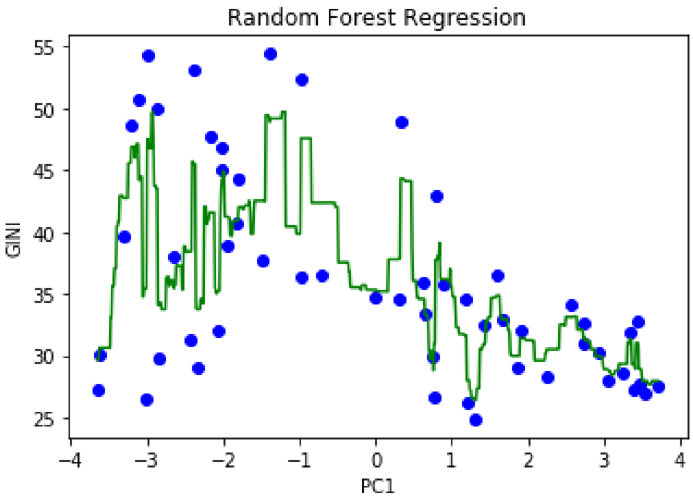
Correlation Table	VA	CoC	GE	PSNV	RQ	RoL
VA	1					
CoC	0,833207	1				
GE	0,809717	0,937356	1			
PSNV	0,752462	0,764514	0,711604	1		
RQ	0,832752	0,885494	0,949644	0,66834	1	
RoL	0,866522	0,961538	0,951996	0,791417	0,919357	1

The table shows the importance of using PCA, since it reduces these six correlated variables into one variable. When we look at the RFR, it gave high R2 value with the training data, but it gave negative R2 value with the test data, which indicates that when the model is fitted for a data, it cannot make good predictions with another data. Briefly, there is a high chance that the model is overfitting the data.

Model 2: In the Model 1, the countries that had even one GINI coefficient were included. For example, the GINI coefficient for Iraq calculated only in 2012 and in the previous data, this value for one year was considered as the average for the 2002-2018 period, which makes the estimations misleading to a certain level. For making the analysis more accurate, the countries that had more than 5 NaN GINI coefficient values for 16 years were eliminated. With the elimination process, 55 countries/territories were gathered. The same model was conducted with the new data and for the training and test split for the RFR, the training data consists of 39 observations and the test data consists of 16 observations. The results are shown below:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.275
Model:                            OLS   Adj. R-squared:                  0.261
Method:                 Least Squares   F-statistic:                     20.06
Date:                Tue, 19 May 2020   Prob (F-statistic):           4.04e-05
Time:                        02:42:25   Log-Likelihood:                -185.83
No. Observations:                  55   AIC:                             375.7
Df Residuals:                      53   BIC:                             379.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     35.9880      0.975     36.911      0.000      34.032      37.944
PC1           -1.8678      0.417     -4.478      0.000      -2.704      -1.031
==============================================================================
Omnibus:                        0.206   Durbin-Watson:                   1.404
Prob(Omnibus):                  0.902   Jarque-Bera (JB):                0.065
Skew:                           0.083   Prob(JB):                        0.968
Kurtosis:                       2.973   Cond. No.                         2.34
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.573
Model:                            OLS   Adj. R-squared:                  0.520
Method:                 Least Squares   F-statistic:                     10.74
Date:                Tue, 19 May 2020   Prob (F-statistic):           1.53e-07
Time:                        02:42:25   Log-Likelihood:                -171.26
No. Observations:                  55   AIC:                             356.5
Df Residuals:                      48   BIC:                             370.6
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     35.8696      1.605     22.348      0.000      32.642      39.097
VA            11.4941      2.579      4.457      0.000       6.309      16.679
CoC            6.8457      3.495      1.959      0.056      -0.180      13.872
GE            -6.2064      5.610     -1.106      0.274     -17.487       5.074
PSNV          -5.5509      2.034     -2.729      0.009      -9.641      -1.461
RQ             0.9946      3.984      0.250      0.804      -7.016       9.005
RoL          -11.6312      6.522     -1.783      0.081     -24.745       1.483
==============================================================================
Omnibus:                        4.463   Durbin-Watson:                   1.956
Prob(Omnibus):                  0.107   Jarque-Bera (JB):                3.706
Skew:                          -0.398   Prob(JB):                        0.157
Kurtosis:                       3.992   Cond. No.                         24.5
==============================================================================
```



With RFR:

Mean Absolute Error (MAE): 7.047063053980342

Mean Squared Error (MSE): 76.09090535597694

Root Mean Squared Error (RMSE): 8.723010108671028

R2 (Training data): 0.8132761503607606

R2 (Test data): -0.3955778175131246

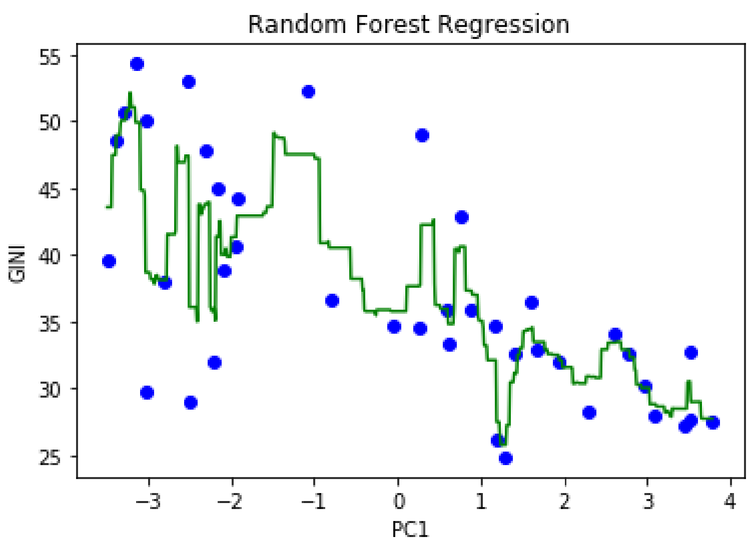
By eliminating the observations with high amount of NaN GINI coefficients, both SLR and MLR models were improved, since they can explain the variability in GINI more. But, when we look at the RFR results, it is visible that MAE, MSE and RMSE were increased, which means that the model makes worse predictions compared to the previous model, and the R2 value for the test data is still negative.

Furthermore, with this new model, the data was divided into training and test data to see whether the model's estimations changed only for the OLS Regression. Again, the data was divided randomly: 15 observations for test and 40 observations for training. The results are shown below:

For training data:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.448
Model:                            OLS   Adj. R-squared:                  0.433
Method:                 Least Squares   F-statistic:                     30.81
Date:                Tue, 19 May 2020   Prob (F-statistic):           2.35e-06
Time:                        05:28:42   Log-Likelihood:                -129.36
No. Observations:                  40   AIC:                             262.7
Df Residuals:                      38   BIC:                             266.1
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     37.1246      0.996     37.261      0.000      35.108      39.142
PC1           -2.3685      0.427     -5.551      0.000      -3.232      -1.505
==============================================================================
Omnibus:                        0.875   Durbin-Watson:                   1.661
Prob(Omnibus):                  0.646   Jarque-Bera (JB):                0.300
Skew:                          -0.181   Prob(JB):                        0.861
Kurtosis:                       3.223   Cond. No.                         2.34
==============================================================================
```

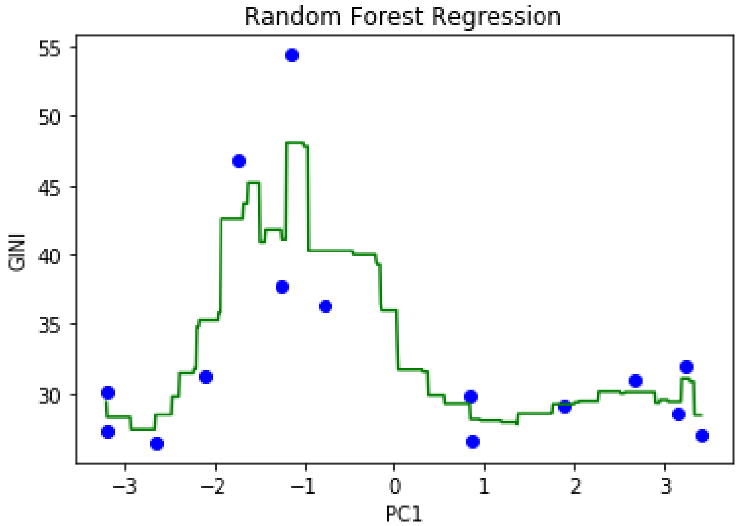
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.700
Model:                            OLS   Adj. R-squared:                  0.646
Method:                 Least Squares   F-statistic:                     12.84
Date:                Tue, 19 May 2020   Prob (F-statistic):           1.96e-07
Time:                        05:28:42   Log-Likelihood:                -117.16
No. Observations:                  40   AIC:                             248.3
Df Residuals:                      33   BIC:                             260.1
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     39.0885      2.101     18.600      0.000      34.813      43.364
VA             9.1532      2.889      3.168      0.003       3.275      15.031
CoC            7.3937      3.823      1.934      0.062      -0.384      15.171
GE           -11.6485      6.376     -1.827      0.077     -24.620       1.323
PSNV          -5.7738      2.031     -2.843      0.008      -9.906      -1.642
RQ             1.6403      4.173      0.393      0.697      -6.849      10.130
RoL           -7.2370      7.715     -0.938      0.355     -22.933       8.459
==============================================================================
Omnibus:                        7.121   Durbin-Watson:                   1.928
Prob(Omnibus):                  0.028   Jarque-Bera (JB):                7.966
Skew:                          -0.486   Prob(JB):                       0.0186
Kurtosis:                       4.958   Cond. No.                         29.3
==============================================================================
```



For test data:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.061
Model:                            OLS   Adj. R-squared:                 -0.011
Method:                 Least Squares   F-statistic:                    0.8457
Date:                Tue, 19 May 2020   Prob (F-statistic):              0.375
Time:                        05:29:15   Log-Likelihood:                -51.500
No. Observations:                  15   AIC:                             107.0
Df Residuals:                      13   BIC:                             108.4
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     32.9573      2.079     15.852      0.000      28.466      37.449
PC1           -0.8146      0.886     -0.920      0.375      -2.728       1.099
==============================================================================
Omnibus:                       10.314   Durbin-Watson:                   2.088
Prob(Omnibus):                  0.006   Jarque-Bera (JB):                6.501
Skew:                           1.422   Prob(JB):                       0.0388
Kurtosis:                       4.521   Cond. No.                         2.35
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.703
Model:                            OLS   Adj. R-squared:                  0.480
Method:                 Least Squares   F-statistic:                     3.155
Date:                Tue, 19 May 2020   Prob (F-statistic):             0.0681
Time:                        05:29:15   Log-Likelihood:                -42.868
No. Observations:                  15   AIC:                             99.74
Df Residuals:                       8   BIC:                             104.7
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     30.1865      2.791     10.817      0.000      23.752      36.622
VA            20.9150      7.611      2.748      0.025       3.364      38.466
CoC           23.9303      8.110      2.951      0.018       5.228      42.632
GE           -24.5766     11.440     -2.148      0.064     -50.957       1.803
PSNV           1.7304      5.120      0.338      0.744     -10.075      13.536
RQ             4.6095     10.183      0.453      0.663     -18.873      28.092
RoL          -24.7409     12.997     -1.904      0.093     -54.711       5.229
==============================================================================
Omnibus:                        0.948   Durbin-Watson:                   2.072
Prob(Omnibus):                  0.623   Jarque-Bera (JB):                0.127
Skew:                           0.208   Prob(JB):                        0.939
Kurtosis:                       3.172   Cond. No.                         26.5
==============================================================================
```



It is visible that although the models with the training data performed much better compared to the previous data, the SLR and MLR models with the test data came out as insignificant because the probability of F-statistic is high. When we look at the SLR with PC1, the

Adjusted R2 is negative, which means that the model cannot explain any variability in the GINI.

Model 3: The model's prediction accuracy on the countries' yearly changing GINI coefficient values was also evaluated. For this purpose, eight countries were pulled from the data, which were:

- Denmark
- Costa Rica
- Moldova
- El Salvador
- Finland
- Honduras
- Indonesia
- Belarus

These are the countries that have less than 2 NaN GINI coefficients. The data for Moldova as an example is shown below:

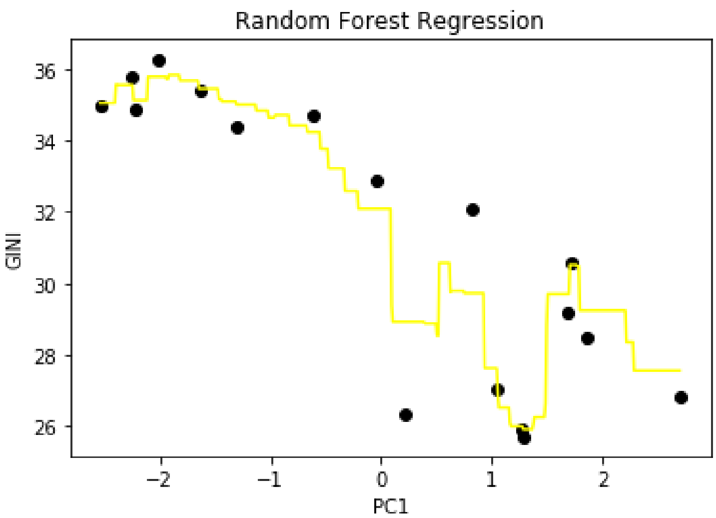
Moldova	VA	CoC	GE	PSNV	RQ	RoL	GINI
2002	-0,53	-0,95	-0,56	-0,16	-0,40	-0,60	35,8
2003	-0,49	-0,85	-0,67	-0,13	-0,45	-0,52	34,9
2004	-0,59	-0,98	-0,87	-0,22	-0,44	-0,33	35
2005	-0,55	-0,67	-0,75	-0,40	-0,46	-0,36	36,3
2006	-0,27	-0,64	-0,79	-0,37	-0,34	-0,52	35,4
2007	-0,29	-0,66	-0,82	-0,02	-0,28	-0,51	34,4
2008	-0,31	-0,63	-0,77	-0,27	-0,18	-0,43	34,7
2009	-0,30	-0,70	-0,56	-0,59	-0,13	-0,44	32,9
2010	-0,06	-0,67	-0,66	-0,38	-0,10	-0,36	32,1
2011	0,05	-0,62	-0,62	-0,05	-0,08	-0,33	30,6
2012	-0,03	-0,61	-0,57	0,05	-0,10	-0,32	29,2
2013	-0,07	-0,75	-0,41	0,00	-0,07	-0,37	28,5
2014	0,01	-0,85	-0,42	-0,16	0,02	-0,25	26,8
2015	0,03	-0,91	-0,65	-0,33	-0,07	-0,35	27
2016	-0,02	-0,95	-0,63	-0,30	-0,11	-0,49	26,3
2017	-0,03	-0,80	-0,53	-0,32	-0,04	-0,41	25,9
2018	-0,11	-0,73	-0,47	-0,35	-0,05	-0,41	25,7

The results for the countries that have too much NaN GINI coefficient values were also calculated by replacing NaN values via zero or multiple imputation. But the results came out as insignificant and that is why, they were not included in the report. The results for these 8 countries are shown below (Note: For RFR, the data was not divided into training and test data, since the number of observations for each country was not enough):

For Moldova:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.715
Model:                            OLS   Adj. R-squared:                  0.696
Method:                 Least Squares   F-statistic:                     37.68
Date:                Thu, 21 May 2020   Prob (F-statistic):           1.90e-05
Time:                        05:49:31   Log-Likelihood:                -36.296
No. Observations:                  17   AIC:                             76.59
Df Residuals:                      15   BIC:                             78.26
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     31.2647      0.528     59.169      0.000      30.138      32.391
PC1           -1.9544      0.318     -6.139      0.000      -2.633      -1.276
==============================================================================
Omnibus:                        2.179   Durbin-Watson:                   0.424
Prob(Omnibus):                  0.336   Jarque-Bera (JB):                1.690
Skew:                          -0.725   Prob(JB):                        0.429
Kurtosis:                       2.465   Cond. No.                         1.66
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.884
Model:                            OLS   Adj. R-squared:                  0.815
Method:                 Least Squares   F-statistic:                     12.73
Date:                Thu, 21 May 2020   Prob (F-statistic):           0.000357
Time:                        05:49:31   Log-Likelihood:                -28.647
No. Observations:                  17   AIC:                             71.29
Df Residuals:                      10   BIC:                             77.13
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     31.3237      4.501      6.959      0.000      21.294      41.353
VA            -9.4981      5.731     -1.657      0.128     -22.267       3.271
CoC            9.7244      3.500      2.778      0.020       1.925      17.523
GE            -5.4338      4.680     -1.161      0.273     -15.861       4.993
PSNV           0.8015      2.772      0.289      0.778      -5.376       6.979
RQ            -6.3760      9.140     -0.698      0.501     -26.742      13.990
RoL           -2.1872      5.471     -0.400      0.698     -14.377      10.003
==============================================================================
Omnibus:                        0.875   Durbin-Watson:                   0.855
Prob(Omnibus):                  0.646   Jarque-Bera (JB):                0.564
Skew:                          -0.424   Prob(JB):                        0.754
Kurtosis:                       2.719   Cond. No.                         41.4
==============================================================================
```



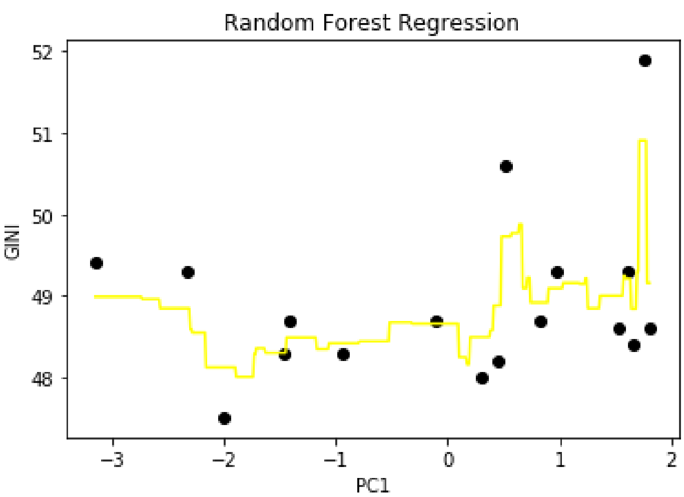
With RFR:

R2: 0.9363575352512211

For Costa Rica:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.063
Model:                            OLS   Adj. R-squared:                  0.001
Method:                 Least Squares   F-statistic:                     1.014
Date:                Thu, 21 May 2020   Prob (F-statistic):              0.330
Time:                        15:56:59   Log-Likelihood:                -23.633
No. Observations:                  17   AIC:                             51.27
Df Residuals:                      15   BIC:                             52.93
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     48.9294      0.251    195.027      0.000      48.395      49.464
PC1           -0.1636      0.162     -1.007      0.330      -0.510       0.183
==============================================================================
Omnibus:                        9.113   Durbin-Watson:                   1.380
Prob(Omnibus):                  0.011   Jarque-Bera (JB):                5.979
Skew:                           1.331   Prob(JB):                       0.0503
Kurtosis:                       4.163   Cond. No.                         1.54
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.779
Model:                            OLS   Adj. R-squared:                  0.647
Method:                 Least Squares   F-statistic:                     5.892
Date:                Thu, 21 May 2020   Prob (F-statistic):            0.00730
Time:                        15:56:59   Log-Likelihood:                -11.339
No. Observations:                  17   AIC:                             36.68
Df Residuals:                      10   BIC:                             42.51
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     54.3565      3.298     16.481      0.000      47.008      61.705
VA            -5.1747      2.851     -1.815      0.100     -11.527       1.178
CoC           -0.3822      1.848     -0.207      0.840      -4.499       3.735
GE             6.0015      2.739      2.191      0.053      -0.101      12.104
PSNV           0.7781      1.462      0.532      0.606      -2.478       4.035
RQ           -13.3979      3.161     -4.238      0.002     -20.442      -6.354
RoL            8.2278      2.946      2.793      0.019       1.664      14.792
==============================================================================
Omnibus:                        0.610   Durbin-Watson:                   2.346
Prob(Omnibus):                  0.737   Jarque-Bera (JB):                0.654
Skew:                          -0.342   Prob(JB):                        0.721
Kurtosis:                       2.325   Cond. No.                         66.7
==============================================================================
```



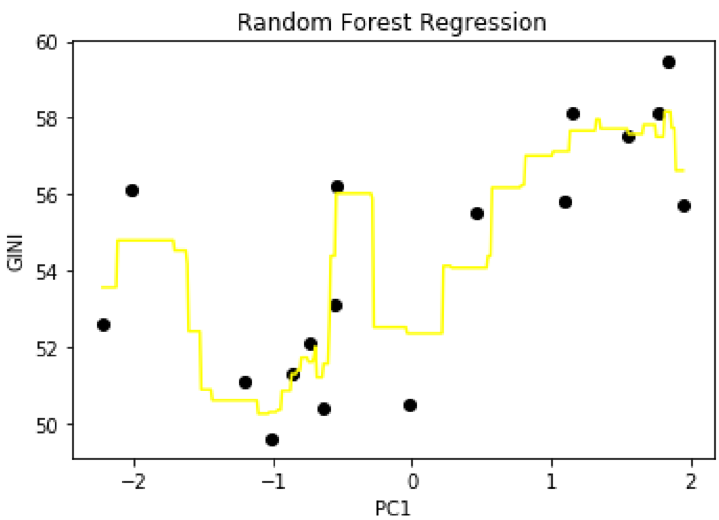
With RFR:

R2: 0.7639789180347719

For Honduras:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.104
Model:                            OLS   Adj. R-squared:                  0.044
Method:                 Least Squares   F-statistic:                     1.742
Date:                Sun, 24 May 2020   Prob (F-statistic):              0.207
Time:                        01:09:03   Log-Likelihood:                -42.181
No. Observations:                  17   AIC:                             88.36
Df Residuals:                      15   BIC:                             90.03
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     54.3059      0.747     72.700      0.000      52.714      55.898
PC1            0.7099      0.538      1.320      0.207      -0.437       1.857
==============================================================================
Omnibus:                        1.385   Durbin-Watson:                   0.643
Prob(Omnibus):                  0.500   Jarque-Bera (JB):                1.128
Skew:                           0.568   Prob(JB):                        0.569
Kurtosis:                       2.449   Cond. No.                         1.39
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.641
Model:                            OLS   Adj. R-squared:                  0.425
Method:                 Least Squares   F-statistic:                     2.973
Date:                Sun, 24 May 2020   Prob (F-statistic):             0.0620
Time:                        01:09:03   Log-Likelihood:                -34.413
No. Observations:                  17   AIC:                             82.83
Df Residuals:                      10   BIC:                             88.66
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     53.1132     11.008      4.825      0.001      28.585      77.641
VA             7.1914      5.787      1.243      0.242      -5.702      20.085
CoC          -13.1264      6.514     -2.015      0.072     -27.640       1.387
GE             6.7166      7.472      0.899      0.390      -9.932      23.365
PSNV          -5.5954      6.401     -0.874      0.403     -19.858       8.667
RQ            -2.0448      6.385     -0.320      0.755     -16.272      12.182
RoL            6.1602      5.698      1.081      0.305      -6.536      18.856
==============================================================================
Omnibus:                        0.916   Durbin-Watson:                   1.610
Prob(Omnibus):                  0.632   Jarque-Bera (JB):                0.703
Skew:                           0.021   Prob(JB):                        0.704
Kurtosis:                       2.005   Cond. No.                         43.3
==============================================================================
```



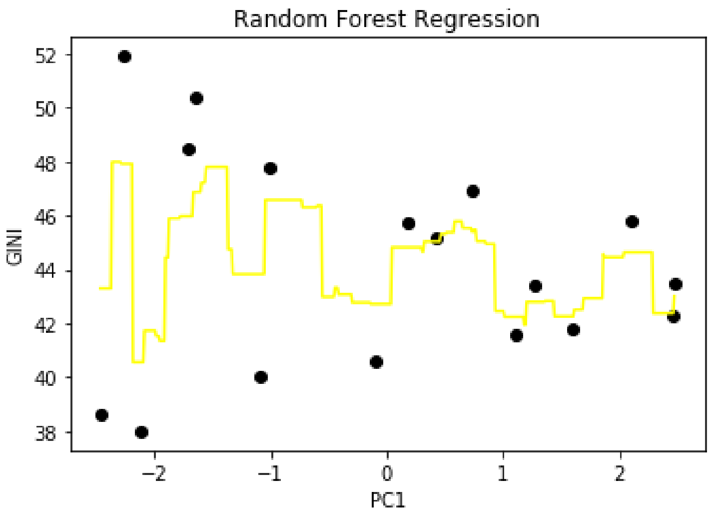
With RFR:

R2: 0.8957517356662641

For El Salvador:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.019
Model:                            OLS   Adj. R-squared:                 -0.047
Method:                 Least Squares   F-statistic:                    0.2840
Date:                Thu, 21 May 2020   Prob (F-statistic):              0.602
Time:                        15:59:40   Log-Likelihood:                -47.143
No. Observations:                  17   AIC:                             98.29
Df Residuals:                      15   BIC:                             99.95
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     44.2353      1.000     44.228      0.000      42.103      46.367
PC1           -0.3225      0.605     -0.533      0.602      -1.612       0.967
==============================================================================
Omnibus:                        0.385   Durbin-Watson:                   0.146
Prob(Omnibus):                  0.825   Jarque-Bera (JB):                0.510
Skew:                          -0.127   Prob(JB):                        0.775
Kurtosis:                       2.190   Cond. No.                         1.65
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.766
Model:                            OLS   Adj. R-squared:                  0.626
Method:                 Least Squares   F-statistic:                     5.456
Date:                Thu, 21 May 2020   Prob (F-statistic):            0.00957
Time:                        15:59:40   Log-Likelihood:                -34.957
No. Observations:                  17   AIC:                             83.91
Df Residuals:                      10   BIC:                             89.75
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     60.4126      4.068     14.851      0.000      51.349      69.476
VA             3.2344      9.208      0.351      0.733     -17.282      23.751
CoC           28.9702     12.067      2.401      0.037       2.083      55.857
GE           -15.4652     12.899     -1.199      0.258     -44.206      13.276
PSNV          18.4582      5.777      3.195      0.010       5.585      31.331
RQ           -11.0111      7.619     -1.445      0.179     -27.986       5.964
RoL            9.4747      5.136      1.845      0.095      -1.970      20.919
==============================================================================
Omnibus:                        1.980   Durbin-Watson:                   1.493
Prob(Omnibus):                  0.371   Jarque-Bera (JB):                1.276
Skew:                          -0.411   Prob(JB):                        0.528
Kurtosis:                       1.938   Cond. No.                         37.4
==============================================================================
```



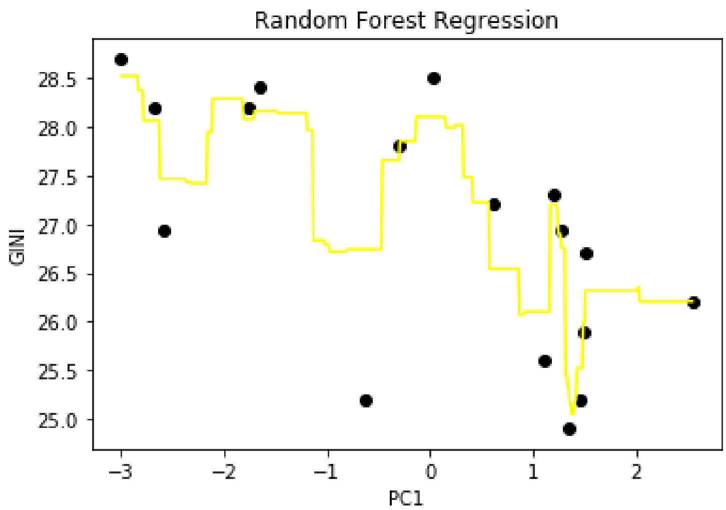
With RFR:

R2: 0.6556232488876609

For Denmark:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.403
Model:                            OLS   Adj. R-squared:                  0.363
Method:                 Least Squares   F-statistic:                     10.13
Date:                Sun, 24 May 2020   Prob (F-statistic):            0.00617
Time:                        01:48:29   Log-Likelihood:                -23.172
No. Observations:                  17   AIC:                             50.34
Df Residuals:                      15   BIC:                             52.01
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     26.9329      0.244    110.306      0.000      26.413      27.453
PC1            0.4589      0.144      3.183      0.006       0.152       0.766
==============================================================================
Omnibus:                        1.494   Durbin-Watson:                   0.998
Prob(Omnibus):                  0.474   Jarque-Bera (JB):                1.239
Skew:                          -0.582   Prob(JB):                        0.538
Kurtosis:                       2.373   Cond. No.                         1.69
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.894
Model:                            OLS   Adj. R-squared:                  0.830
Method:                 Least Squares   F-statistic:                     14.03
Date:                Sun, 24 May 2020   Prob (F-statistic):           0.000235
Time:                        01:48:29   Log-Likelihood:                -8.4956
No. Observations:                  17   AIC:                             30.99
Df Residuals:                      10   BIC:                             36.82
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     34.7788      6.326      5.498      0.000      20.683      48.874
VA            -3.3587      1.841     -1.824      0.098      -7.461       0.744
CoC            7.2231      2.931      2.464      0.033       0.692      13.755
GE            -9.3097      1.512     -6.157      0.000     -12.679      -5.941
PSNV          -1.0846      0.988     -1.098      0.298      -3.286       1.117
RQ             1.9718      2.294      0.860      0.410      -3.140       7.083
RoL           -1.3265      2.058     -0.645      0.534      -5.912       3.259
==============================================================================
Omnibus:                        0.506   Durbin-Watson:                   3.084
Prob(Omnibus):                  0.776   Jarque-Bera (JB):                0.486
Skew:                          -0.339   Prob(JB):                        0.784
Kurtosis:                       2.523   Cond. No.                         246.
==============================================================================
```



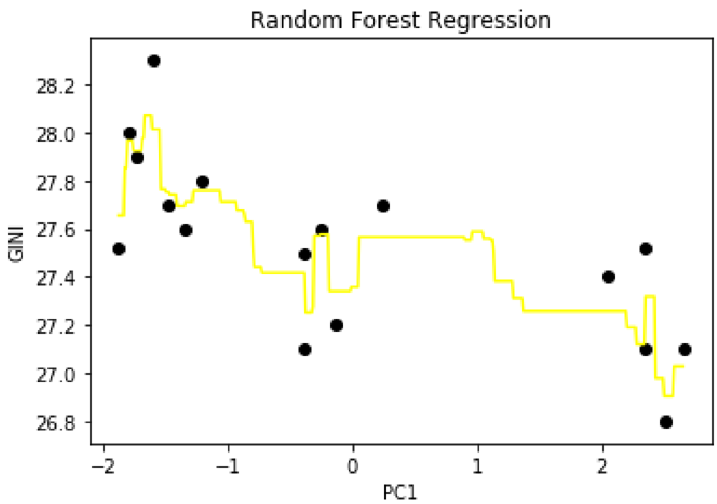
With RFR:

R2: 0.8393830230898719

For Finland:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.508
Model:                            OLS   Adj. R-squared:                  0.476
Method:                 Least Squares   F-statistic:                     15.51
Date:                Sun, 24 May 2020   Prob (F-statistic):            0.00132
Time:                        03:35:54   Log-Likelihood:                -1.0996
No. Observations:                  17   AIC:                             6.199
Df Residuals:                      15   BIC:                             7.866
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     27.5200      0.067    412.895      0.000      27.378      27.662
PC1           -0.1583      0.040     -3.938      0.001      -0.244      -0.073
==============================================================================
Omnibus:                        0.012   Durbin-Watson:                   1.050
Prob(Omnibus):                  0.994   Jarque-Bera (JB):                0.194
Skew:                           0.048   Prob(JB):                        0.907
Kurtosis:                       2.485   Cond. No.                         1.66
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.730
Model:                            OLS   Adj. R-squared:                  0.568
Method:                 Least Squares   F-statistic:                     4.502
Date:                Sun, 24 May 2020   Prob (F-statistic):             0.0182
Time:                        03:35:54   Log-Likelihood:                 3.9897
No. Observations:                  17   AIC:                             6.021
Df Residuals:                      10   BIC:                             11.85
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     28.4230      4.528      6.277      0.000      18.333      38.513
VA            -0.2970      0.806     -0.368      0.720      -2.093       1.499
CoC            1.7244      1.070      1.612      0.138      -0.659       4.108
GE             1.0265      0.815      1.260      0.236      -0.789       2.842
PSNV          -0.4747      0.625     -0.759      0.465      -1.868       0.919
RQ            -1.6445      0.872     -1.886      0.089      -3.588       0.298
RoL           -1.4930      1.907     -0.783      0.452      -5.741       2.755
==============================================================================
Omnibus:                        0.272   Durbin-Watson:                   1.845
Prob(Omnibus):                  0.873   Jarque-Bera (JB):                0.164
Skew:                           0.198   Prob(JB):                        0.921
Kurtosis:                       2.726   Cond. No.                         371.
==============================================================================
```



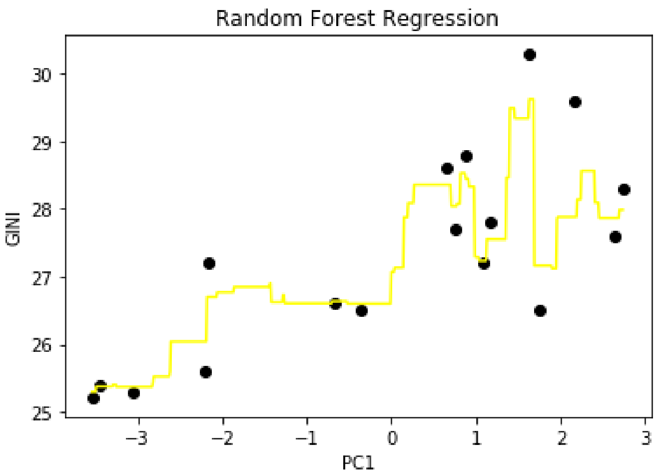
With RFR:

R2: 0.8860474336793529

For Belarus:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.594
Model:                            OLS   Adj. R-squared:                  0.567
Method:                 Least Squares   F-statistic:                     21.94
Date:                Sun, 24 May 2020   Prob (F-statistic):           0.000294
Time:                        03:40:03   Log-Likelihood:                -22.939
No. Observations:                  17   AIC:                             49.88
Df Residuals:                      15   BIC:                             51.55
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     27.3059      0.241    113.373      0.000      26.793      27.819
PC1           -0.5436      0.116     -4.684      0.000      -0.791      -0.296
==============================================================================
Omnibus:                        1.046   Durbin-Watson:                   1.497
Prob(Omnibus):                  0.593   Jarque-Bera (JB):                0.580
Skew:                           0.444   Prob(JB):                        0.748
Kurtosis:                       2.820   Cond. No.                         2.08
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.791
Model:                            OLS   Adj. R-squared:                  0.665
Method:                 Least Squares   F-statistic:                     6.295
Date:                Sun, 24 May 2020   Prob (F-statistic):            0.00576
Time:                        03:40:04   Log-Likelihood:                -17.308
No. Observations:                  17   AIC:                             48.62
Df Residuals:                      10   BIC:                             54.45
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     28.4563      3.928      7.244      0.000      19.704      37.208
VA             6.4485      3.239      1.991      0.074      -0.768      13.665
CoC            2.9438      4.724      0.623      0.547      -7.582      13.469
GE            -3.4709      2.621     -1.324      0.215      -9.312       2.370
PSNV           0.0239      1.128      0.021      0.983      -2.490       2.538
RQ            -3.8506      2.123     -1.814      0.100      -8.580       0.879
RoL           -2.7999      3.513     -0.797      0.444     -10.628       5.028
==============================================================================
Omnibus:                        0.213   Durbin-Watson:                   2.376
Prob(Omnibus):                  0.899   Jarque-Bera (JB):                0.091
Skew:                           0.135   Prob(JB):                        0.955
Kurtosis:                       2.764   Cond. No.                         74.1
==============================================================================
```



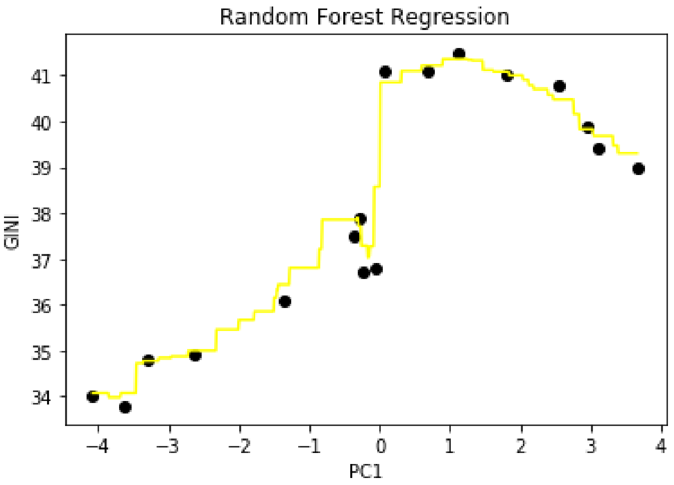
For RFR:

R2: 0.8669009602872324

For Indonesia:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.687
Model:                            OLS   Adj. R-squared:                  0.666
Method:                 Least Squares   F-statistic:                     32.89
Date:                Sun, 24 May 2020   Prob (F-statistic):           3.94e-05
Time:                        03:46:18   Log-Likelihood:                -30.531
No. Observations:                  17   AIC:                             65.06
Df Residuals:                      15   BIC:                             66.73
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     38.0176      0.376    100.996      0.000      37.215      38.820
PC1           -0.9220      0.161     -5.735      0.000      -1.265      -0.579
==============================================================================
Omnibus:                        2.152   Durbin-Watson:                   0.628
Prob(Omnibus):                  0.341   Jarque-Bera (JB):                1.546
Skew:                           0.715   Prob(JB):                        0.462
Kurtosis:                       2.628   Cond. No.                         2.34
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   GINI   R-squared:                       0.881
Model:                            OLS   Adj. R-squared:                  0.809
Method:                 Least Squares   F-statistic:                     12.30
Date:                Sun, 24 May 2020   Prob (F-statistic):           0.000413
Time:                        03:46:18   Log-Likelihood:                -22.331
No. Observations:                  17   AIC:                             58.66
Df Residuals:                      10   BIC:                             64.50
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     52.7728      5.150     10.247      0.000      41.298      64.247
VA           -17.8753     11.641     -1.536      0.156     -43.812       8.061
CoC            5.8457      4.190      1.395      0.193      -3.490      15.181
GE           -15.4243      5.769     -2.674      0.023     -28.278      -2.571
PSNV           4.8652      2.066      2.355      0.040       0.262       9.468
RQ             6.5171      6.852      0.951      0.364      -8.750      21.784
RoL           12.5390      7.145      1.755      0.110      -3.381      28.459
==============================================================================
Omnibus:                        3.303   Durbin-Watson:                   1.515
Prob(Omnibus):                  0.192   Jarque-Bera (JB):                1.418
Skew:                          -0.637   Prob(JB):                        0.492
Kurtosis:                       3.615   Cond. No.                         92.3
==============================================================================
```



For RFR:

R2: 0.9632667874410626

The results from each country gave mixed findings. From the regressions, it is visible that the models for Moldova, Belarus, Indonesia, Finland, and Denmark were significant, especially for the first three countries, since their Adjusted R2 values from SLR with PC1 were 50% or higher. But, when we look at their probability of F-statistic from SLR with PC1, the models for Costa Rica, El Salvador, and Honduras were insignificant, especially for El Salvador, which its probability of F-statistic was 60.2%. This issue was tried to be understood by looking at the data for El Salvador:

El Salvador	VA	CoC	GE	PSNV	RQ	RoL	GINI
2002	0,12	-0,68	-0,50	0,30	0,02	-0,51	51,9
2003	0,11	-0,37	-0,35	-0,18	-0,14	-0,46	50,4
2004	0,02	-0,42	-0,34	-0,07	0,06	-0,38	47,8
2005	-0,08	-0,45	-0,38	-0,03	-0,18	-0,45	48,5
2006	0,10	-0,30	-0,15	-0,17	0,04	-0,62	45,7
2007	0,09	-0,37	-0,20	-0,01	0,19	-0,65	45,2
2008	0,11	-0,36	-0,16	0,04	0,21	-0,67	46,9
2009	0,06	-0,25	-0,04	-0,01	0,32	-0,75	45,8
2010	0,07	-0,28	0,01	0,06	0,36	-0,81	43,5
2011	0,04	-0,26	-0,10	0,11	0,47	-0,72	42,3
2012	-0,04	-0,41	-0,13	0,23	0,33	-0,70	41,8
2013	-0,01	-0,35	-0,12	-0,03	0,31	-0,62	43,4
2014	0,15	-0,38	-0,03	-0,02	0,32	-0,48	41,6
2015	0,15	-0,42	-0,24	-0,02	0,20	-0,59	40,6
2016	0,19	-0,52	-0,29	-0,10	0,09	-0,71	40
2017	0,15	-0,51	-0,36	-0,26	-0,15	-0,86	38
2018	0,04	-0,59	-0,45	-0,33	-0,04	-0,82	38,6

When we look at the GINI coefficient for El Salvador, it is visible that it gradually decreased from 51.9 in 2002 to 38.6 in 2018, which is a positive progress in terms of reducing inequality. But when we look at the six WGI variables for El Salvador, it is also observable that the values for the current years are worse than the previous ones. This shows that the values for GINI coefficient and six WGI variables are not exactly negatively correlated, at least for El Salvador.

CONCLUSION

When the average of seven variables was taken for each country and regression models were conducted (ex. Model 1 and Model 2), the models of SLR with PC1 and MLR with the six WGI variables could not explain the majority of variability in the GINI coefficient but came out as significant. But, when the models were calculated with the test data, which was done for RFR in both Model 1 and Model 2 and for SLR and MLR in the Model 2, the models came out as insignificant with negative Adjusted R2 values. Furthermore, when the values of seven variables were taken from the year 2002 to 2018 for several countries and the models were conducted for each country individually (ex. Model 3), the models came out as significant and highly elucidator for many, but not for all the countries. For that, it can be concluded that enough evidence was not found for the assumptions and hypothesis indicated in the paper.

REFERENCES

1) WGI 2019 Interactive Home - World Bank Group. (2019). Retrieved May 24, 2020, from https://info.worldbank.org/governance/wgi/
2) Chappelow, J. (2020, January 29). Gini Index Definition. Retrieved May 24, 2020, from https://www.investopedia.com/terms/g/gini-index.asp