The Labor Force Participation Rate in the Context of ESG Models at World Level

Leogrande, Angelo and Costantiello, Alberto

LUM University Giuseppe Degennaro, LUM Enterprise s.r.l.

1 June 2023

Angelo Leogrande[1]°, Alberto Costantiello[2]**
*LUM University Giuseppe Degennaro, Casamassima, Bari, Puglia, Italy, EU
°LUM Enterprise s.r.l., Casamassima, Bari, Puglia, Italy, EU

The Labor Force Participation Rate in the Context of ESG Models at World Level

Abstract

In this article we analyze the impact of Labor Force Partecipation Rate-LFPR in the context of the Environmental, Social and Governance-ESG model at world level. We use data from the ESG dataset of the World Bank for the period 2011-2020. We use Panel Data with Fixed Effects, Panel Data with Random Effects, Pooled OLS, Dynamic Panel. We find that the level of LFPR is positively associated among others to "*Ratio of Female to Male Labor Force Participation Rate*" and "*Life Expectancy at Birth*", and negatively associated among others, to "*Unemployment*" and "*Agricultural Land*". Furthermore, we have applied a clusterization with the k-Means algorithm optimized with the Silhouette coefficient, and we found the presence of three clusters. Finally, we confront eight different machine learning algorithms to predict the value of LFPR. We find that the best predictor is the Linear Regression. Linear Regression predicts an increase in LFPR equal to 0.42% on average for the analyzed countries.

Keywords*: Analysis of Collective Decision-Making, General, Political Processes: Rent-Seeking, Lobbying, Elections, Legislatures, and Voting Behaviour, Bureaucracy, Administrative Processes in Public Organizations, Corruption, Positive Analysis of Policy Formulation, Implementation.*

JEL Classification: D7, D70, D72, D73, D78.

1) Introduction-Research Question

In the following article, we analyze the role of LFPR in the context of ESG models worldwide using data of the World Bank. The role of work in the current economic systems is subjected to many pressures and criticisms. On the one hand, in fact, technology and capital tend to reduce the role of labour by favoring IT systems and the work of machines. On the other hand, labour seems to be increasingly a privilege rather than a right in developed and underdeveloped economies due to gender and racial discrimination. However, work remains an essential component for value added both in countries with low per capita incomes and in countries with high per capita incomes. The value of labour, and in particular the value of LFPR tends to reflect all the various discriminations present at the social level. Added to these discriminations are those produced by educational qualifications with the contrast between high-skilled workers and low-skilled workers. Labour is therefore the most

[1]Professor of Economics at LUM University Giuseppe Degennaro and Researcher at LUM Enterprise s.r.l. Email: leogrande.cultore@lum.it, Strada Statale 100 km 18, Casamassima, Bari, Puglia, Italia.
[2]Professor of Economics at LUM University Giuseppe Degennaro. Email: costantiello@lum.it. Strada Statale 100 km 18, Casamassima, Bari, Puglia, Italia.

relevant context for analyzing the social and economic contradictions in terms of income inequality and opportunities.
In our case, we analyzed the LFPR variable in the context of ESG models. This analysis adds an element of originality to the study. In fact, if the various determinants of the labor market have been extensively analyzed in the scientific literature, there are few studies that have taken into consideration the relationship between the LFPR variable and ESG models.
The article continues as follows: the second section refers to the analysis of the literature, the third section presents the econometric model, the fourth section shows the results of the clustering, the fifth section presents the results of the analysis with the machine algorithms learning, the sixth section concludes.

2) Literature Review

A brief review of the literature related to LFPR is presented below. The articles discussed are not exhaustive of the scientific debate in terms of LFPR. The citations reported have the sole purpose of introducing the topic by highlighting the salient points of recent research.

LFPR, Covid 19 and health issues. Low levels of LFPR have been positively associated to high levels of child abuse and neglect during the Covid 19 pandemic at Los Angeles [1]. The increase in drug abuse during the Covid 19 pandemics has reduced the level of LFPR in the post-pandemic [2]. There is a negative relationship between depression and anxiety disorders on one side and LFPR on the other side [3]. The presence of health and psychological problems reduces the population's ability to actively participate in the labor market. The distinction between people who enjoy good health and people who have both physical and mental health problems creates a further discrimination in the workplace. Furthermore, the fact that people actively participate in the labor market also reduces the likelihood that they are involved in criminal activity or violent behavior. High levels of LFPR are therefore necessary either to ensure a better health status of the population either to reduce the incidence of domestic and social violence.

LFPR and developing countries. Remittances reduces the female LFPR but do not affect male LFPR in a set of 122 developing countries in the period 1990-2015 [4]. There is a negative relationship between the level of female education and the female LFPR in rural India [5]. A study on individuals in the age 55-64 in Turkey shows that the main determinants for LFPR are education and marital status [6]. The increase of LFPR has a positive effect on income distribution in India [7]. There is a positive relationship between female LFPR and female salaries in Pakistan [8]. There is a negative relationship between female LFPR and natality in Brazil [9]. Female LFPR and occupational segregation grown together in 66 developing countries in the period 1980-2011 [10]. Female LFPR in Middle East is low despite the increase in the level of female education; the presence of a male's veto power on the decision of woman to work reduces female LFPR [11]. The analysis shows some counterfactual elements. The fact that there is an inverse relationship between female education and female LFPR levels in India may be due to the low incomes in rural markets especially for high-skilled female workers. But, in general, except for this case, the dynamics of LFPR and especially female LFPR in developing and low-income countries is similar to that of high-income countries.

LFPR and democracy and discriminations. There is a positive relationship between female LFPR and the participation in USA democratic elections i.e. the number of voters increases with female LFPR [12]. The black-white racial gap in LFPR is greater than the hispanic-white gap LFPR in USA [13]. The presence of low skilled immigrants has a positive effect on LFPR for low skilled coloured native born workers in South Africa [14]. Labor participation rates have a positive impact in terms of democratic activism. A person who works might also have a greater interest in voting at political elections. There is therefore a positive relationship between the value of participation in the labor market and the value of participation in democratic life, at national level. However, for the same reasons, racial discriminations could have a negative retroactive effect on the solidity of democratic institutions. In fact, the case of the USA shows that both the Afro-American community and the Hispanic community, despite showing a proclivity for activism in the labor market, still have reduced levels of employability, due to racial discrimination.

LFPR policies. The offering of afterschool care increases labor force participation rate by 7% in Chile [15]. The promotion of childcare services has increased mother's LFPR by 0.2% in Germany, especially for mothers with medium-high education [16]. The investment in labor-intensive industries can improve female LFPR in Botswana and Namibia [17]. There are policies that can have a positive effect on the growth of the value of LFPR, especially in the case of female LFPR. Specifically, it is possible to develop care services for children, to ensure that women can actively participate in the labor market. However, in some cases, such as in the case of people with low levels of education, the offer of childcare services should also be accompanied by a set of educational and financial interventions to support mothers in the pathways of entry into the labor market. Economic policies that invest in labor-intensive sectors can help LFPR growth.

LFPR miscellaneous. There is a positive relationship between the increase in female entrepreneurship and the reduction in male vs. female LFPR [18]. Internet has increased female LFPR by 4.1% for married woman thanks to tele working [19]. A positive relationship was found between labour supply and LFPR in Germany during the period 2003-2010 [20]. The promotion of female entrepreneurship can reduce the gap between men and women in the sense of LFPR. However, it is necessary to invest in promoting women's economic rights, freedom, and capabilities to increase the value of LFPR through female entrepreneurship. One of the variables that increases women's participation in work is information technology. In fact, if IT makes it possible to improve the work-life balance, then it is possible that the number of women participating in the labor market will increase. In addition, there are demographic conditions that can lead to an increase in the value of LFPR. However, it is said that an increase in the value of LFPR is not always accompanied by an improvement in the condition of the workers in the sense of income and the quality of workplace.

3) The Econometric Model for the Estimation of the Value of LFPR

Below we present a regression analysis aimed at identifying the LFPR determinants within the context of the ESG dataset of the World Bank. The data refer to 193 countries in the period between 2011 and 2020. The data were analyzed using the following econometric models or: Panel Dat with Fixed Effects, Panel Data with Random Effects, Pooled OLS, Dynamic Panel. In summary, we estimated the following equation or:

$$\boldsymbol{LFPR_{it} = \alpha_i + \beta_1(AL)_{it} + \beta_2(CO_2)_{it} + \beta_3(LEAB)_{it} + \beta_4(PM2.5)_{it} + \beta_5(RFTM)_{it} + \beta_6(RQ)_{it} + \beta_7(UT)_{it}}$$

Where $\boldsymbol{i = 193}$ and $\boldsymbol{t = [2011; 2020]}$.

We found that LFPR is positively associated to:

- CO_2: it is a variable that considers the value of carbon dioxide emissions deriving from the combustion of fossil fuels and from the production of cement. They include carbon dioxide produced during the burning of solid, liquid, and gaseous fuels and gas flaring. There is a positive relationship between the value of LFPR and CO_2. This positive relationship is due to the fact that the countries where there is the greatest production of CO_2 are also the countries that have the highest levels of LFPR, i.e. the western countries. Countries that have high levels of LFPR are also countries that have an active industrial and manufacturing system, and which therefore tend to have significant CO_2 production. However, this positive relationship is very likely to change in the future due to the emphasis placed by European governments especially on climate change. The incentives that are offered to the European industrial system try to transform production through the application of sustainable methodologies that could allow an increase in the LFPR in the reduction of the CO_2 value in the future [21].

- *LEAB:* is a variable that considers life expectancy at birth as the number of years a newborn would live if the prevailing patterns of mortality at the time of birth remained the same throughout life. There is a positive relationship between the LEAB value and the LFPR value. Such a relationship tends to be paradoxical. Countries that have high levels of LFPR are either low-income countries from Central and Southern Africa, either high-income countries i.e. North America, North Europe and Oceania. Otherwise, there are a set of countries that show a low level of LFPR i.e. Mediterranean and South Asia countries. However, the fact that there is a positive relationship between LFPR and LEAB suggests that the positive effect on life expectancy in high-income countries that have high levels of LFPR tends to largely offset the negative effect on life expectancy in countries lower-middle-income with a high level of LFPR.
- *PM2.5:* is a variable that considers the weighted exposure for the population to environmental pollution from PM2.5 defined as the average level of exposure of the population of a nation to concentrations of suspended particles measuring less than 2.5 microns in aerodynamic diameter, which are able to penetrate deep into the respiratory tract and cause great harm to health. harm. Exposure is calculated by weighting the average annual concentrations of PM2.5 by population in urban and rural areas. There is a positive relationship between the PM2.5 value and the LFPR value. In fact, it must be considered that many of the countries that have high levels of LFPR are low-middle income countries that also have high levels of pollution from PM2.5 emissions. In fact, countries in Central and Southern Africa that have high levels of LFPR also have high levels of PM2.5.
- *RFTM:* is a variable that considers the percentage of the population aged 15 or over who is economically active: all the people who provide work to produce goods and services during a given period. The ratio of female to male labor force participation rates is calculated by dividing the female labor force participation rate by the male labor force participation rate and multiplying by 100. Looking at the map of countries by RFTM value many countries that have high LFPR levels also have high RFTM values. This is the case, for example, of the countries of Central-Southern Africa, of North America, of Northern Europe, of South America and of some parts of Oceania. Hence the positive relationship between the LFPR value and the RFTM value derives. However, it must also be considered that there are some countries which have low LFPR value, and which nonetheless have high RFTM values such as Italy, Spain, China for example.
- *RQ:* captures perceptions of government's ability to formulate and implement robust policies and regulations that enable and promote private sector development. The estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5. There is a positive relationship between the RQ value and the LFPR value. This relationship is since many of the countries that have high levels of LFPR such as the countries of North America and Northern Europe also have high levels of RQ. However, there are also countries that have high LFPR lifelines and low RQ levels such as, for example, the countries of Central-Southern Africa. It therefore follows that the positive impact of high-middle-income per capita countries with high LFPR levels on the RQ value tends to more than offset the negative impact of low-middle-income countries with high LFPR levels on the RQ value [22].

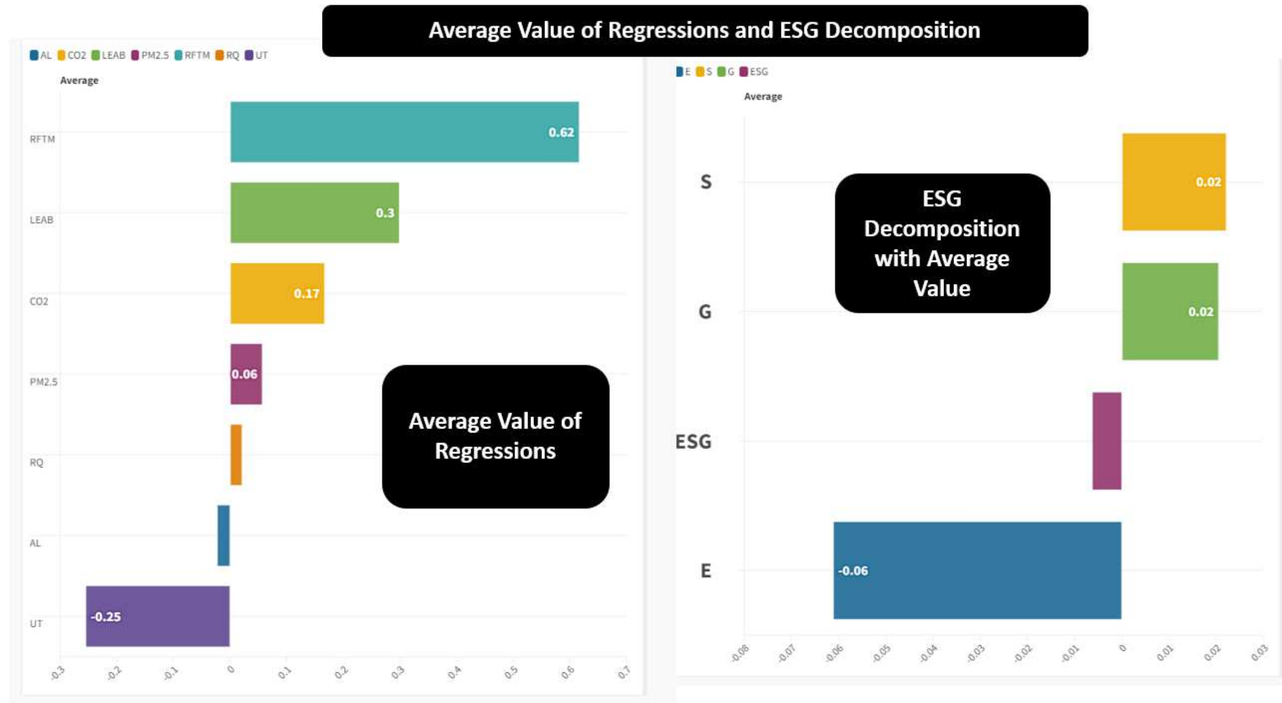


Figure 1. Average value of regressions and ESG decomposition with average value.

LFPR is negatively associated to:

- *AL:* refers to the share of arable land, with permanent crops and permanent pastures. Arable land includes land defined by FAO as land under temporary crops (double harvested areas are counted once), temporary grassland for mowing or grazing, land under market or market gardens and land temporarily fallow. There is a positive relationship between the AL value and the LFPR value. In fact, many countries that have high levels of AL have low levels of LFPR such as for example India and many countries in the Mediterranean area. On the contrary, many Anglo-Saxon countries that have a high level of LFPR, such as the USA, Canada, Australia have low levels of AL.
- *UT:* refers to the share of the labor force that is out of work but available and looking for work. There is a negative relationship between the UT value and the LFPR value. That is, unemployment tends to rise as the value of LFPR decreases. This is because many countries that have high LFPR values are also low unemployment countries such as North America, Northern Europe, and Australia.

		ACRONYM	Pooled OLS		Dynamic Panel		Random-Effects		Fixed-Effects		Average
			Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	
	const		2,33246	***	-0,138629	*	3,47774	***	3,61957	***	2,32279
E	A5	AL	-0,0388912	***	-0,0179322	***	-0,016645	***	-0,0145975	**	-0,022
E	A11	CO2	0,22287	***	0,172915	***	0,135394	***	0,133879	***	0,16626
S	A34	LEAB	0,289417	***	0,308939	***	0,298095	***	0,296225	***	0,29817
E	A46	PM2.5	0,127636	***	0,0129159	***	0,043316	***	0,040214	***	0,05602
E	A54	RFTM	0,59414	***	0,624443	***	0,621311	***	0,623847	***	0,61594
G	A55	RQ	0,435197	***	-0,20831	**	-0,0559749	**	-0,0892834	***	0,02041
S	A65	UT	-0,0992825	***	-0,248144	**	-0,324306	***	-0,344686	***	-0,2541
	A33(-1)				0,42001						

4) Clusterization with k-Means Algorithm Optimized with the Silhouette Coefficient

In the following analysis we apply the k-Means clustering algorithm to check for clustering within the data. Since the k-Means algorithm is an unsupervised algorithm then we will apply the Silhouette coefficient to choose the optimal number of clusters. The result shows the presence of three clusters. Considering the value of the median of the clusters, the following ordering of the clusters results: $C3 = 69{,}978 > C1 = 59{,}79 > C2 = 43{,}677$.

- *Cluster 1*: Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Belgium, Burkina Faso, Bangladesh, Bulgaria, Belarus, Belize, Brazil, Barbados, Brunei Darussalam, Bhutan, Botswana, Canada, Channel Islands, Chile, Cote d'Ivoire, Congo Dem. Rep., Cabo Verde, Costa Rica, Cuba, Cyprus, Czechia, Germany, Denmark, Dominican Republic, Ecuador, Spain, Estonia, Finland, Fiji, France, United Kingdom, Georgia, Guinea, The Gambia, Guinea Bissau, Equatorial Guinea, Guatemala, Guam, Guyana, Hong Kong SAR, Honduras, Haiti, Hungary, Ireland, Israel, Jamaica, Japan, Kyrgyz Republic, South Korea, Lao PDR, Sri Lanka, Lesotho, Lithuania, Luxembourg, Latvia, Maldives, Mexico, North Macedonia, Malta, Myanmar, Montenegro, Mongolia, Mauritius, Malaysia, Namibia, New Caledonia, Nigeria, Nicaragua, Netherlands, Norway, Panama, Philippines, Poland, Portugal, French Polynesia, Romania, Russian Federation, Rwanda, Saudi Arabia, Sierra Leone, El Salvador, Sao Tome and Principe, Suriname, Slovak Republic, Slovenia, Sweden, Chad, Togo, Tonga, Trinidad and Tobago, Ukraine, Uruguay, United States, Uzbekistan, St. Vincent and the Grenadines, Venezuela, Virgin Islands, Samoa, South Africa, Zambia, Zimbabwe. It is the second cluster for median value of LFPR. It is a very large cluster made up of various countries that are either upper-middle income, such as the USA, Australia, Austria, Hong Kong, either lower-middle income countries such as Mongolia, Namibia, Zimbabwe. We can therefore note that although these countries have the same capacity in terms of LFPR value, they nevertheless have important differences in terms of GDP value both in absolute value and in per capita value. This analysis highlights the low value of labor compared to capital in creating the conditions for economic growth. In fact, even if the population actively participates in the labor market, this does not necessarily lead to growth in GDP. Otherwise, it is the capital endowment that allows, LFPR ceteris paribus, to create the conditions for economic growth either in the sense of growth in the value of per capita gross domestic product either as growth in GDP in absolute terms. Furthermore, we must also underline the enormous heterogeneity of the conditions of workers in the various countries even if they are in the same cluster for the median level of LFPR. In fact, working class conditions in Mongolia are certainly not comparable to those in Sweden, despite Mongolia and Sweden participate in the same cluster in the sense of LFPR. And in fact, perhaps it would be necessary to create a new indicator capable of measuring LFPR adjusted for the level of workers conditions.

Figure 2. Structure of the clusters with the k-Means algorithm optimized with the Silhouette Coefficient.

- *Cluster 2:* Afghanistan, Bosnia and Herzegovina, Comoros, Djibouti, Algeria, Egypt Arab Rep., Gabon, Greece, Croatia, India, Iran, Iraq, Italy, Jordan, Lebanon, Libya, Morocco, Moldova, Mauritania, Nepal, Pakistan, Papua New Guinea, Puerto Rico, West Bank and Gaza, Sudan, Senegal, Somalia, Serbia, Eswatini, Syrian Arab Republic, Tajikistan, Turkmenistan, Tunisia, Turkey, Yemen. It is the last cluster by value of LFPR. C2 is a very heterogeneous cluster as it is made up of countries from various continents, which have different levels of per capita income. All these countries have low participation in the labor market. However, it must be considered that the LFPR variable could underestimate the presence of undeclared, irregular, and informal workers in the economy. In Italy, for example, there is a significant percentage of the population working in the irregular and in the informal economy. These are workers who however appear as inactive in the official labour statistics. It is probable that similar conditions also occur in other countries of C2. However, net of this effect, it is highly probable that there is a problem of inefficient incentives in these countries. The workers of the C2 countries do not find in the labor market a feasible opportunity to improve their life and their social condition through the active participation in the job market. One of the reasons for the low participation of workers in the labor market could be the low level of labor income. Other reasons could be connected to the lack of development of industrial and labor systems capable of offering jobs considered financially and technically adequate. These countries suffer for a double problem in the labor market: on the one hand, low wages and on the other, insufficient working conditions. C2 is the only cluster that does not include high-income countries. This condition indicates that low levels of LFPR could in the long run compromise per capita income and productivity levels at country level.
- *Cluster 3:* Angola, United Arab Emirates, Burundi, Benin, Bahrain, The Bahamas, Bolivia, Central African Republic, Switzerland, China, Cameroon, Congo Rep., Colombia, Eritrea, Ethiopia, Ghana, Indonesia, Iceland, Kazakhstan, Kenya, Cambodia, Kuwait, Liberia, St. Lucia, Macao SAR, Madagascar, Mali, Mozambique, Malawi, Niger, New Zealand, Oman, Peru, North Korea, Paraguay, Qatar, Singapore, Solomon Islands, South Sudan, Thailand, Timor-Leste, Tanzania, Uganda, Vietnam, Vanuatu. This is the first cluster in the sense of

LFPR. The cluster is made up of countries with low-medium per capita income, except for the following countries: Switzerland, Iceland, New Zealand, Singapore, Qatar, Macao, United Arab Emirates. The other countries that are in the cluster, all have low per capita incomes. It must be considered that many of these countries are very significant nations from the point of view of absolute GDP value as in the case of China and Indonesia. Other countries have a reduced GDP value both from a per capita and absolute perspective. Many of the C3 countries are African and Latin American ones. It must therefore be considered that there is a real dichotomy considering the relationship between the LFPR and the value of GDP either in absolute value either per capita. Indeed, the LFPR tends to be high both in countries with high levels of per capita income and in countries with low levels of per capita income. The C3 therefore presents a very significant level of polarization between rich and poor countries. Both are characterized by the presence of high labor participation of the population. This condition makes us understand how much work is in some ways irrelevant. In fact, the real distinction between rich and poor countries, as indicated in the case of C3, does not consist in the difference in terms of LFPR, but in the capital endowment. This fact is so evident that even with the same values of LFPR there are still significant differences in terms of per capita income, which are entirely attributable to differences in capital endowments.

5) Machine Learning and Prediction for the Estimation of the Future Value of LFPR

Below we present an analysis for predicting the future value of LFPR. Specifically, we compare eight different machine learning algorithms. The algorithms are compared according to their ability to maximize the R-squared value and to minimize the MAE-Mean Average Error, MSE-Mean Standard Error, RMSE-Root Mean Standard Error value. The algorithms were trained with 70% of the available data while the remaining 30% was used for the actual prediction. To identify the best performing algorithm, four rankings were created for each of the four statistical indicators presented. The score of each algorithm in the individual rankings was added up. The algorithm with the lowest aggregate score was chosen, i.e. the algorithm that totaled the highest places in the single rankings presented. In this way, the following ordering of the algorithms was obtained, i.e.:

- Linear Regression with a payoff equal to 5;
- Random Forest Regression with a payoff value of 7;
- Tree Ensemble Regression with a payoff value of 13;
- Gradient Boosted Tree Regression with a payoff value of 15;
- PNN-Probabilistic Neural Network with a payoff value of 20;
- ANN-Artificial Neural Network with a payoff value of 24;
- Simple Regression Tree with a payoff value of 29;
- Polynomial Regression with a payoff value of 31.

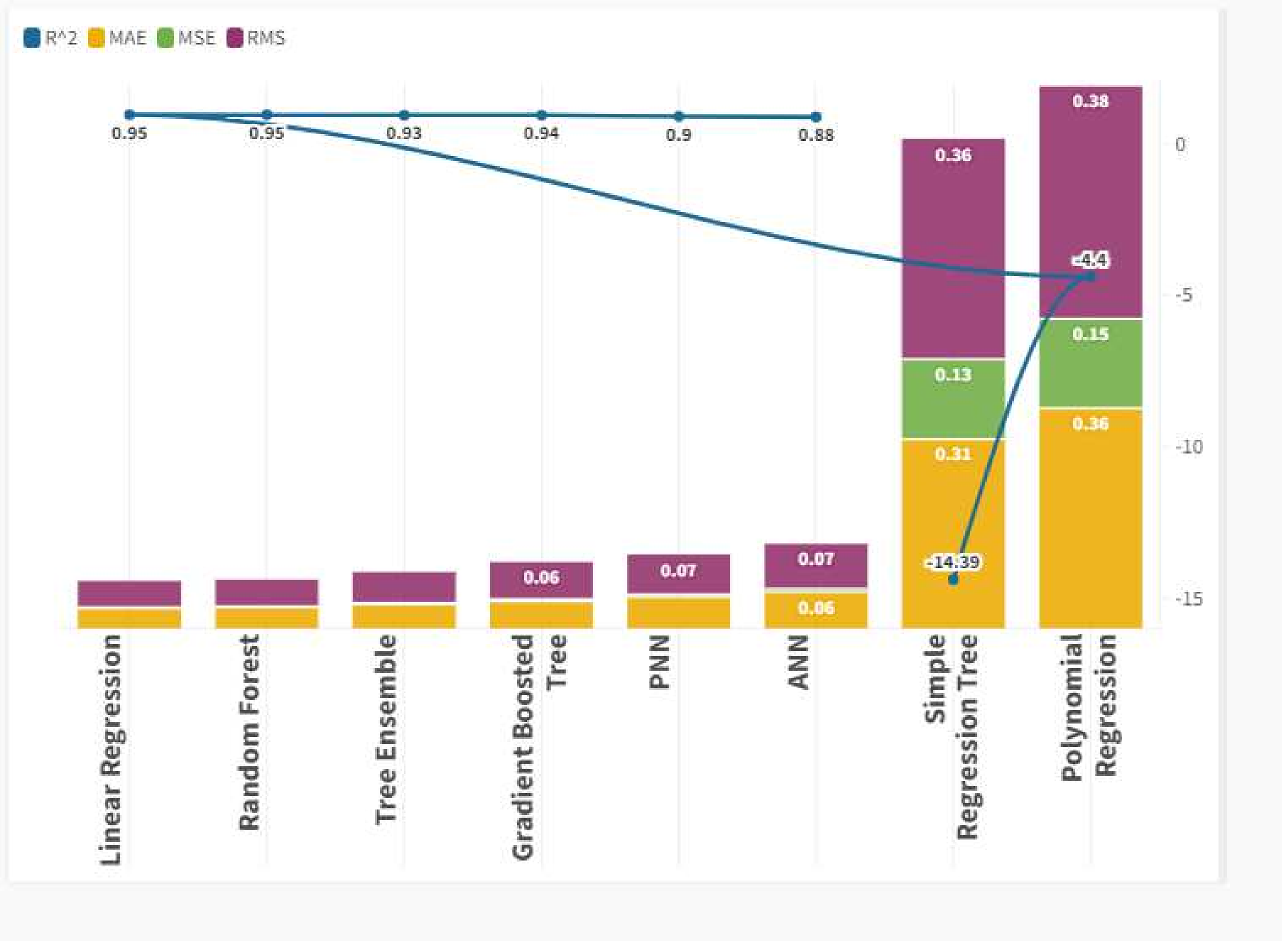


Figure 3. Ranking of algorithms based on the maximization of R-squared and Minimization of MAE, MSE and RMSE.

Therefore, by applying the best predictor algorithm or Linear Regression it is possible to verify that there are some countries for which an increase in the LFPR value is expected, while there are other countries for which a reduction is predicted. There are some countries for which a growth in the value of LFPR is predicted, namely: Philippines with a value of +9.1%; Venezuela with +7.88%; Ecuador with 6.61%; Lebanon with 5.9%; Cuba with 5.67%; Honduras with a value of 5.58%; El Salvador with a value of 5.24%; Uruguay with a value of 3.95%; The Bahamas with a value of 3.48%; Bolivia with 3.26%; Mali with 3.21%; Trinidad and Tobago with 2.53%; Gabon with 2.33%; Iceland with 2.24%; Kuwait with 2.05%, Bhutan with 1.96%,Vietnam with 1.91%; Canada with 0.86%; Belize with 0.82%; Equatorial Guinea with 0.81%; Algeria with 0.5%; Germany 0.34%; United States '0.11%; North Korea with 0.08%; Ukraine with 0.06%.

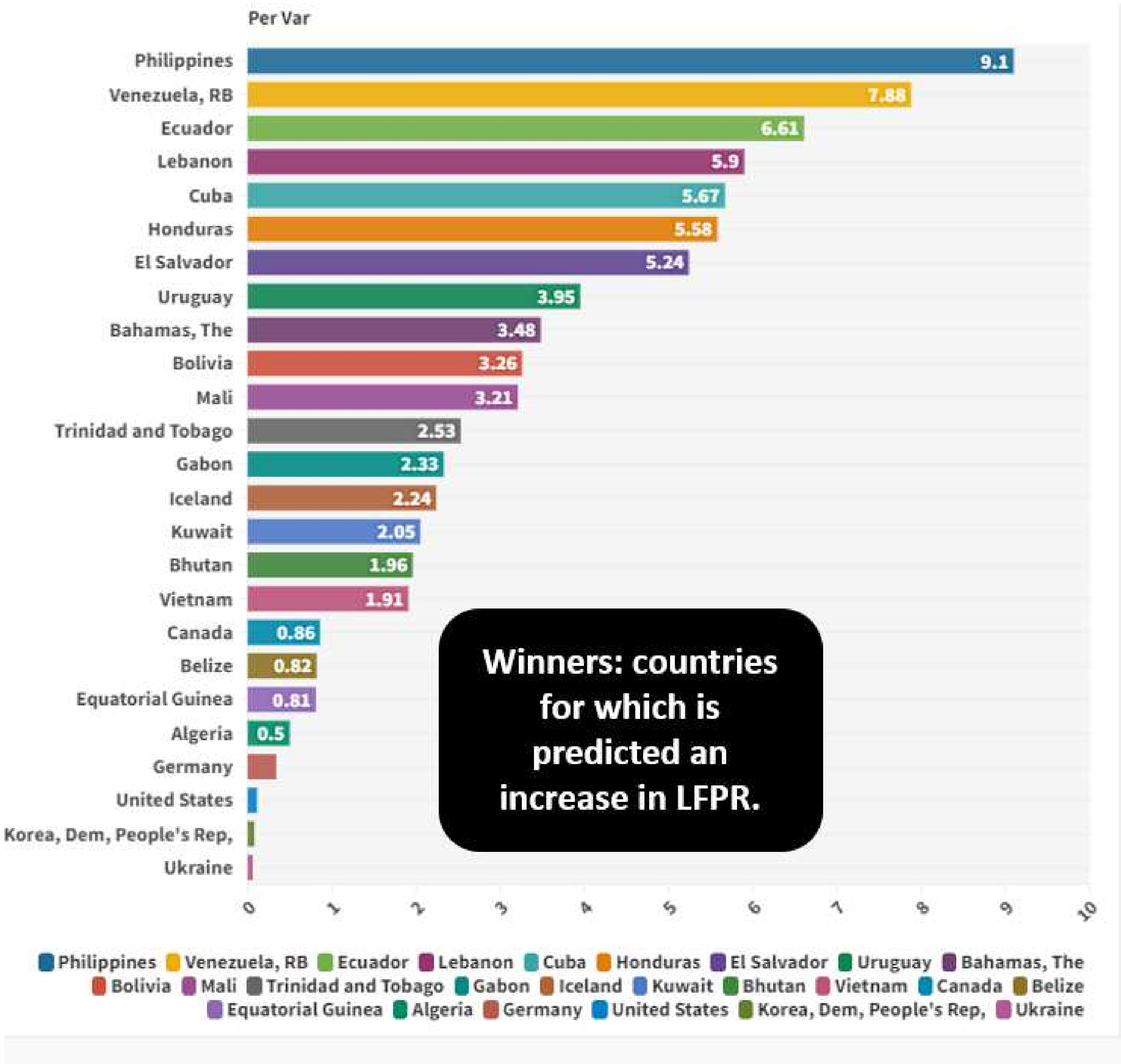


Figure 4. Winners: countries for which is predicted an increase in LFPR.

By applying the best predictor algorithm or Linear Regression it is also possible to obtain the following values for the countries that are losers or: Nicaragua with -0.07%; Guyana with -0.14%; Sri Lanka with -0.22%; Madagascar with -0.24%; Kazakhstan with -0.28%; Mozambique with -0.32%; Greece with -0.36%; Congo Dem Rep with -0.4%; Ghana with -0.47%; Lesotho with -0.59%; Myanmar with -0.61%; Vanuatu with -0.78%; Guam with -0.95%; Guinea-Bissau with -1.11%; French Polynesia with -1.18%; United Kingdom with -1.36%; Denmark with -1.37%; Norway with -1.39%; Mauritania with -1.45%; Virgin Islands with -1.5%; Georgia with -1.53%; Belgium with -1.71%; Poland with -1.78%; Malaysia with -1.96%; Croatia with -1.97%; Netherlands with -2.13%; Slovenia with -2.14%; Hungary with -2.77%; Zambia with -4.16%; Malta with -5.09%; Azerbaijan with -5.26%; Jordan with -6.13%.

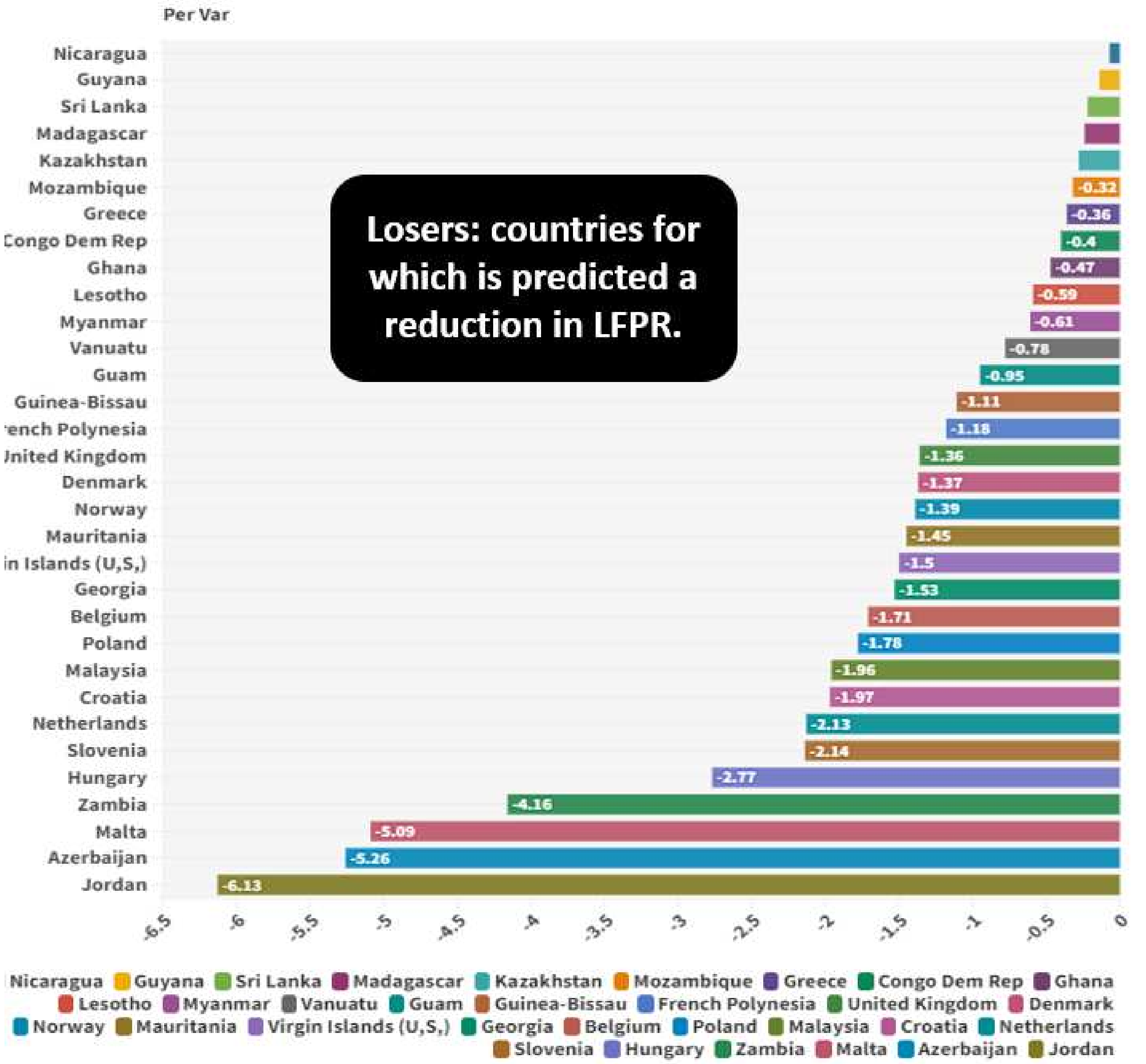


Figure 15. Losers. Countries for which is predicted a reduction in LFPR.

6) Conclusions

In this article, we analyzed the relationship between LFPR and ESG using a dataset from the World Bank over the period 2011-2020. We found that the value of LFPR is negatively connected to the E-Environment component, and positively connected to the S-Social and G-Governance components within the ESG model. Thus, econometric analysis shows that LFPR growth tends to be costly for the Environment even if it is compatible with Social and Governance issues within the ESG model.
The cluster analysis performed with the k-Means algorithm optimized with the Silhouette Coefficient shows the existence of three clusters in the sense of LFPR. The k-Means algorithm shows that within the same cluster it is possible to find either countries with low per capita income either countries with high per capita income. LFPR growth by itself does not increase GDP. The positive relationship between LFPR and GDP per capita is mediated through the investment in capital structure. Therefore, to ensure that the increase in the LFPR in low-income countries has a significant positive impact in terms of GDP, it is necessary to invest in capital endowments in developing countries.
Finally, the predictive analysis based on machine learning algorithms shows a positive trend in the value of the LFPR in the future with an average growth of 0.42% for the countries analyzed.

7) Declarations

Data Availability Statement. The data presented in this study are available on request from the corresponding author.

Funding. The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Competing Interest. The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication.

Software. The authors have used the following software: Gretl for the econometric models, Orange for clusterization and network analysis, and KNIME for machine learning and predictions. They are all free version without licenses.

Acknowledgements. We are grateful to the teaching staff of the LUM University "Giuseppe Degennaro" and to the management of the LUM Enterprise s.r.l. for the constant inspiration to continue our scientific research work undeterred.

8) Bibliography

[1] G. E. Barboza, L. B. Schiamberg and L. Pachl, "A spatiotemporal analysis of the impact of COVID-19 on child abuse and neglect in the city of Los Angeles, California," *Child Abuse & Neglect,* vol. 116, no. 104740, 2021.

[2] J. Greenwood, N. Guner and K. Kopecky, "Did Substance Abuse During the Pandemic Reduce Labor Force Participation?," *Federal Reserve Bank of Atlanta Policy Hub,* 2022.

[3] S. Banerjee, P. Chatterji and K. Lahiri, "Effects of psychiatric disorders on labor market outcomes: a latent variable approach using multiple clinical indicators," *Health economics,* vol. 26, no. 2, pp. 184-205, 2017.

[4] S. Azizi, "The impacts of workers' remittances on human capital and labor supply in developing countries," *Economic Modelling,* vol. 75, pp. 377-396, 2018.

[5] F. Afridi, T. Dinkelman and K. Mahajan, "Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades," *Journal of Population Economics,* vol. 31, pp. 783-818, 2018.

[6] İ. ŞENTÜRK, "Aging and Employment: A Study on the Labor Force Participation Decision of the Elderly in Turkey," *Bulletin of Business and Economics (BBE),* vol. 10, no. 4, pp. 211-215, 2021.

[7] A. C. Angeles, S. J. Donaire, R. B. Reyes and R. Cabauatan, "The Effect of Gender Inequality in Education, Labor Force Participation and Economic Opportunity on the Income Distribution of India," *Jurnal Studi Guru dan Pembelajaran,* vol. 4, no. 3, pp. 781-794, 2021.

[8] J. R. Afridi, M. Khan, H. U. Khan, L. Muhammad and Z. Iqbal, "Female Labor Force Participation in Pakistan and Central Asian States: A Comparative Analysis," *International Journal of Disaster Recovery and Business Community,* vol. 12, no. 1, pp. 560-68, 2021.

[9] L. R. De Souza, E. L. Rios-Neto and B. L. Queiroz, "The Changing Effects of Fertility on Women's Education and Labor Supply: Natural Experiments at Different Parities in Brazil," 2019.

[10] M. Borrowman and S. Klasen, "Drivers of gendered sectoral and occupational segregation in developing countries," *Feminist Economics,* vol. 26, no. 2, pp. 62-94, 2020.

[11] M. .. Majbouri, "Preferences and the Puzzle of Female Labor Force Participation in the Middle East," *ERF 29th Annual Conference,* 2023.

[12] R. J. Cebula and G. M. Alexander, "Female Labor Force Participation and Voter Turnout: Evidence from the American Presidential Elections," *Review of Economics and Institutions,* vol. 2, no. 29, p. 8, 2017.

[13] T. Cajner, T. Radler, D. Ratner and I. Vidangos, "Racial gaps in labor market outcomes in the last four decades and over the business cycle," *Finance and Economics Discussion Series 2017-071. Washington: Board of Governors of the Federal Reserve System,* 2017.

[14] N. H. Broussard, "Immigration and the labor market outcomes of natives in developing countries: A case study of South Africa," *Economic Development and Cultural Change ,* vol. 65, no. 3, pp. 389-424, 2017.

[15] C. Martínez and M. Perticará, "Childcare effects on maternal employment: Evidence from Chile," *Journal of Development Economics,* vol. 126, pp. 127-137, 2017.

[16] K. U. Müller and K. Wrohlich, "Does subsidized care for toddlers increase maternal labor supply. Evidence from a large-scale expansion of early childcare," *DIW Discussion Papers,* no. 1747, 2018.

[17] M. A. Matandare, "Botswana unemployment rate trends by gender: Relative analysis with upper middle income Southern African countries (2000-2016)," *Dutch Journal of Finance and Management,* vol. 2(2), no. 04, 2018.

[18] G. Ribes-Giner, I. Moya-Clemente, R. Cervelló-Royo and M. R. Perello-Marin, "Domestic economic and social conditions empowering female entrepreneurship," *Journal of Business Research,* vol. 89, pp. 182-189, 2018.

[19] L. J. Dettling, "Broadband in the labor market: The impact of residential high-speed internet on married women's labor force participation," *Ilr Review,* vol. 70, no. 2, pp. 451-482, 2017.

[20] M. C. Burda and S. Seele, "Reevaluating the German labor market miracle," *German Economic Review,* vol. 21, no. 2, pp. 139-179., 2020.

[21] A. Costantiello and A. Leogrande, "The Determinants of CO2 Emissions in the Context of ESG Models at World Level," *University Library of Munich, Germany.,* no. 117110, 2023.

[22] A. Costantiello and A. Leogrande, "The Regulatory Quality and ESG Model at World Level," *University Library of Munich, Germany,* no. 116663, 2023.

9) Appendix

ESG	LABEL	ACRONYM	VARIABLE	DESCRIPTION
S	A33	LFPR	Labor force participation rate	Labor force participation rate is the proportion of the population ages 15 and older that is economically active: all people who supply labor for the production of goods and services during a specified period.
E	A5	AL	Agricultural land (% of land area)	Agricultural land refers to the share of land area that is arable, under permanent crops, and under permanent pastures. Arable land includes land defined by the FAO as land under temporary crops (double-cropped areas are counted once), temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land temporarily fallow. Land abandoned as a result of shifting cultivation is excluded. Land under permanent crops is land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest, such as cocoa, coffee, and rubber. This category includes land under flowering shrubs, fruit trees, nut trees, and vines, but excludes land under trees grown for wood or timber. Permanent pasture is land used for five or more years for forage, including natural and cultivated crops.
E	A11	CO2	CO_2 emissions (metric tons per capita)	Carbon dioxide emissions are those stemming from the burning of fossil fuels and the manufacture of cement. They include carbon dioxide produced during consumption of solid, liquid, and gas fuels and gas flaring.
S	A34	LEAB	Life expectancy at birth, total (years)	Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

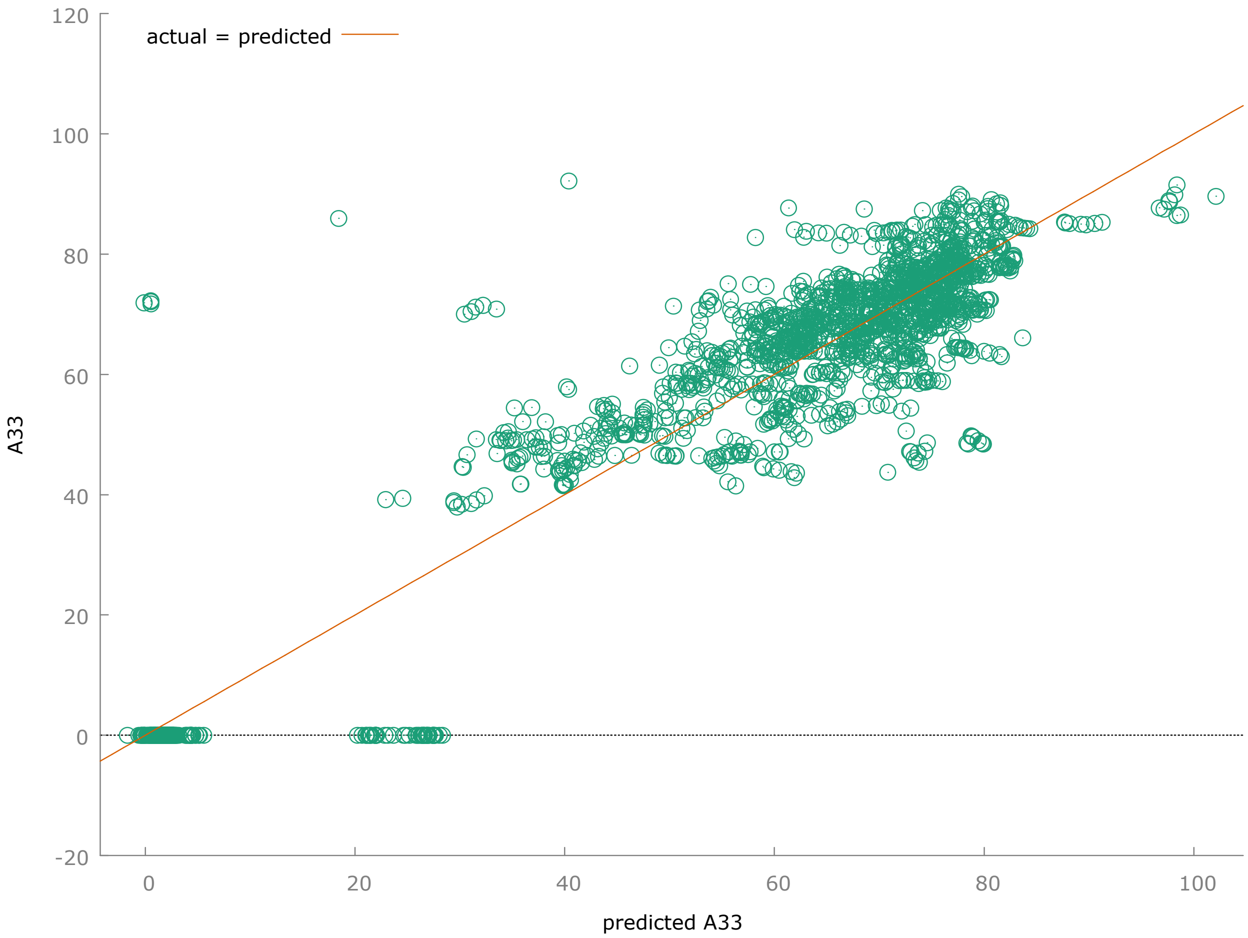
E	A46	PM2.5	PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)	Population-weighted exposure to ambient PM2.5 pollution is defined as the average level of exposure of a nation's population to concentrations of suspended particles measuring less than 2.5 microns in aerodynamic diameter, which are capable of penetrating deep into the respiratory tract and causing severe health damage. Exposure is calculated by weighting mean annual concentrations of PM2.5 by population in both urban and rural areas.
E	A54	RFTM	Ratio of female to male labor force participation rate (%) (modeled ILO estimate)	Labor force participation rate is the proportion of the population ages 15 and older that is economically active: all people who supply labor for the production of goods and services during a specified period. Ratio of female to male labor force participation rate is calculated by dividing female labor force participation rate by male labor force participation rate and multiplying by 100.
G	A55	RQ	Regulatory Quality: Estimate	Regulatory Quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.
S	A65	UT	Unemployment, total (% of total labor force) (modeled ILO estimate)	Unemployment refers to the share of the labor force that is without work but available for and seeking employment.

Pooled OLS, using 1930 observations
Included 193 cross-sectional units
Time-series length = 10
Dependent variable: A33

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	2,33246	0,578591	4,031	<0,0001	***
A5	−0,0388912	0,00957361	−4,062	<0,0001	***
A11	0,222870	0,0478485	4,658	<0,0001	***
A34	0,289417	0,0120673	23,98	<0,0001	***
A46	0,127636	0,0113117	11,28	<0,0001	***
A54	0,594140	0,00929476	63,92	<0,0001	***
A55	0,435197	0,0238527	18,25	<0,0001	***
A65	−0,0992825	0,0366400	−2,710	0,0068	***

Mean dependent var	56,37973	S.D. dependent var	27,26489
Sum squared resid	164775,4	S.E. of regression	9,259117
R-squared	0,885091	Adjusted R-squared	0,884673

F(7, 1922)	2114,904	P-value(F)	0,000000
Log-likelihood	−7029,968	Akaike criterion	14075,94
Schwarz criterion	14120,46	Hannan-Quinn	14092,31
rho	0,856063	Durbin-Watson	0,229849



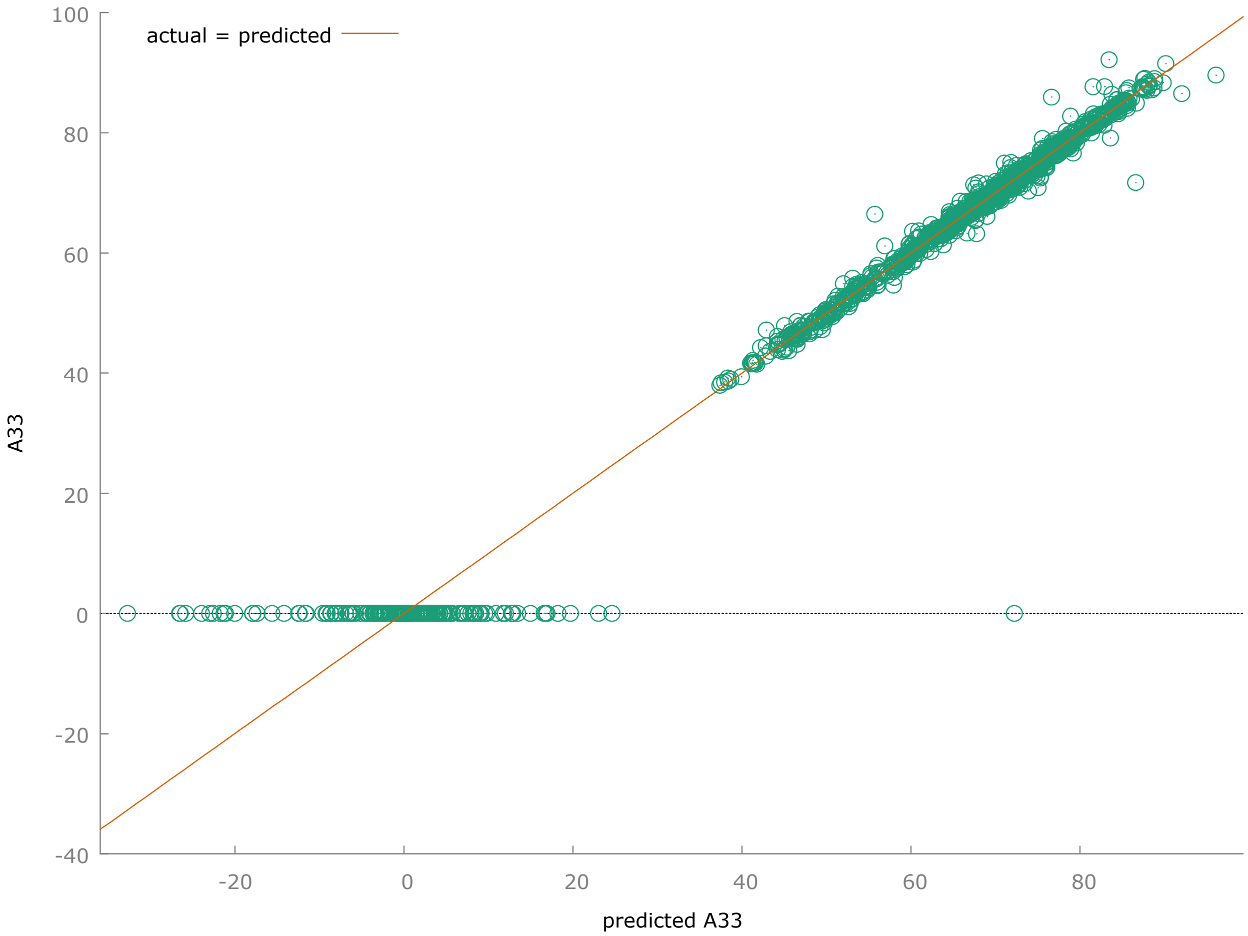
1-step dynamic panel, using 1544 observations
Included 193 cross-sectional units
H-matrix as per Ox/DPD
Dependent variable: A33

	Coefficient	*Std. Error*	*z*	*p-value*	
A33(-1)	0,420010	0,435579	0,9643	0,3349	
const	−0,138629	0,0756895	−1,832	0,0670	*
A5	−0,0179322	0,00464803	−3,858	0,0001	***
A11	0,172915	0,0570130	3,033	0,0024	***
A34	0,308939	0,0491474	6,286	<0,0001	***
A46	0,0129159	0,00383798	3,365	0,0008	***
A54	0,624443	0,0469425	13,30	<0,0001	***
A55	−0,208310	0,0811506	−2,567	0,0103	**

A65	−0,248144	0,125547	−1,977	0,0481	**

Sum squared resid	22746,16	S.E. of regression	3,849461

Number of instruments = 44
Test for AR(1) errors: z = -1,63743 [0,1015]
Test for AR(2) errors: z = 0,0789429 [0,9371]
Sargan over-identification test: Chi-square(35) = 1363,95 [0,0000]
Wald (joint) test: Chi-square(8) = 17245,1 [0,0000]



Random-effects (GLS), using 1930 observations
Using Nerlove's transformation
Included 193 cross-sectional units
Time-series length = 10
Dependent variable: A33

	Coefficient	Std. Error	z	p-value	
const	3,47774	0,891186	3,902	<0,0001	***
A5	−0,0166450	0,00634004	−2,625	0,0087	***
A11	0,135394	0,0367374	3,685	0,0002	***
A34	0,298095	0,0105128	28,36	<0,0001	***
A46	0,0433160	0,00665693	6,507	<0,0001	***
A54	0,621311	0,0101810	61,03	<0,0001	***
A55	−0,0559749	0,0220200	−2,542	0,0110	**
A65	−0,324306	0,0514203	−6,307	<0,0001	***

Mean dependent var	56,37973	S.D. dependent var	27,26489
Sum squared resid	215394,2	S.E. of regression	10,58345
Log-likelihood	−7288,478	Akaike criterion	14592,96
Schwarz criterion	14637,48	Hannan-Quinn	14609,33
rho	0,286115	Durbin-Watson	1,064062

'Between' variance = 104,011

'Within' variance = 11,7229

theta used for quasi-demeaning = 0,894429

Joint test on named regressors -

Asymptotic test statistic: Chi-square(7) = 56540,9

with p-value = 0

Breusch-Pagan test -

Null hypothesis: Variance of the unit-specific error = 0

Asymptotic test statistic: Chi-square(1) = 5219,67

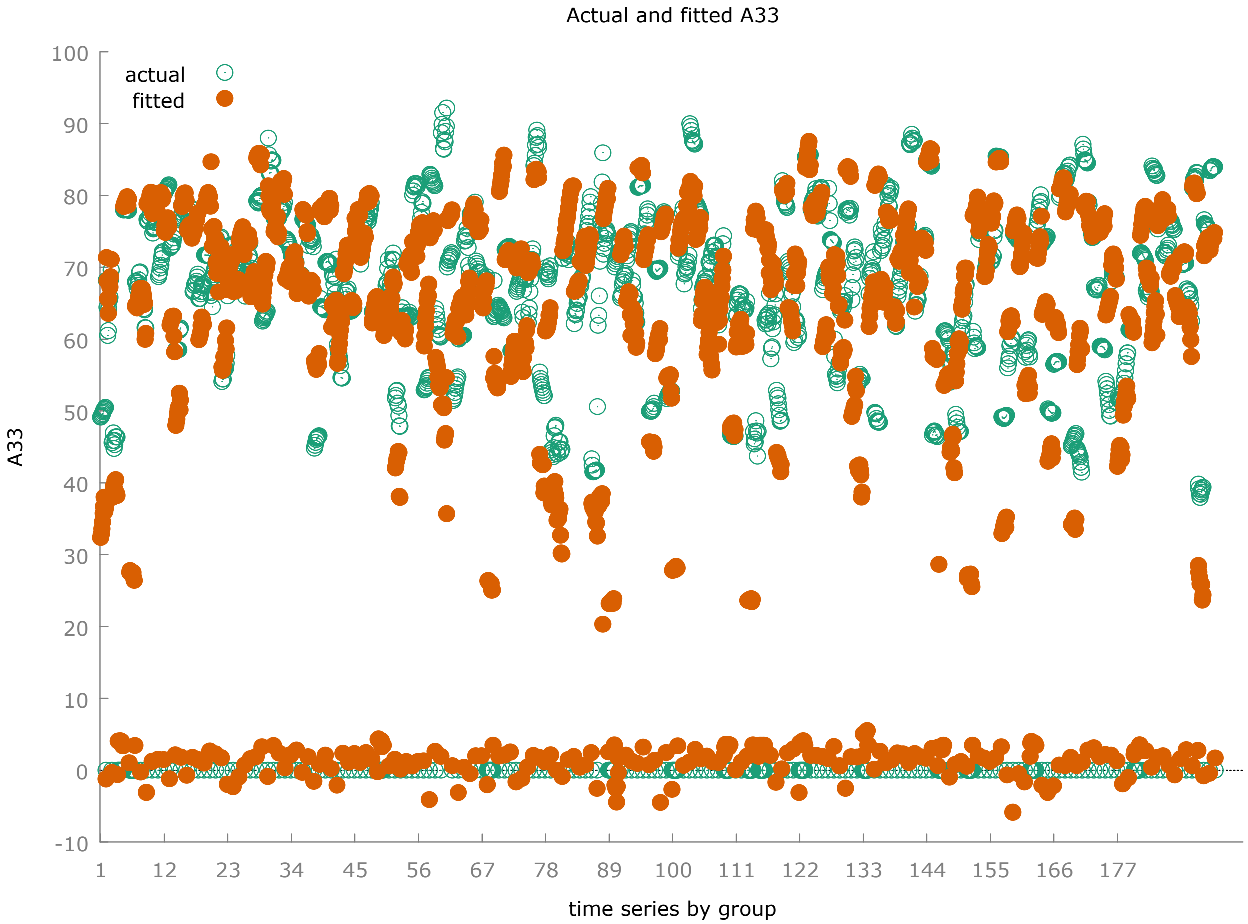
with p-value = 0

Hausman test -

Null hypothesis: GLS estimates are consistent

Asymptotic test statistic: Chi-square(7) = 61,3738

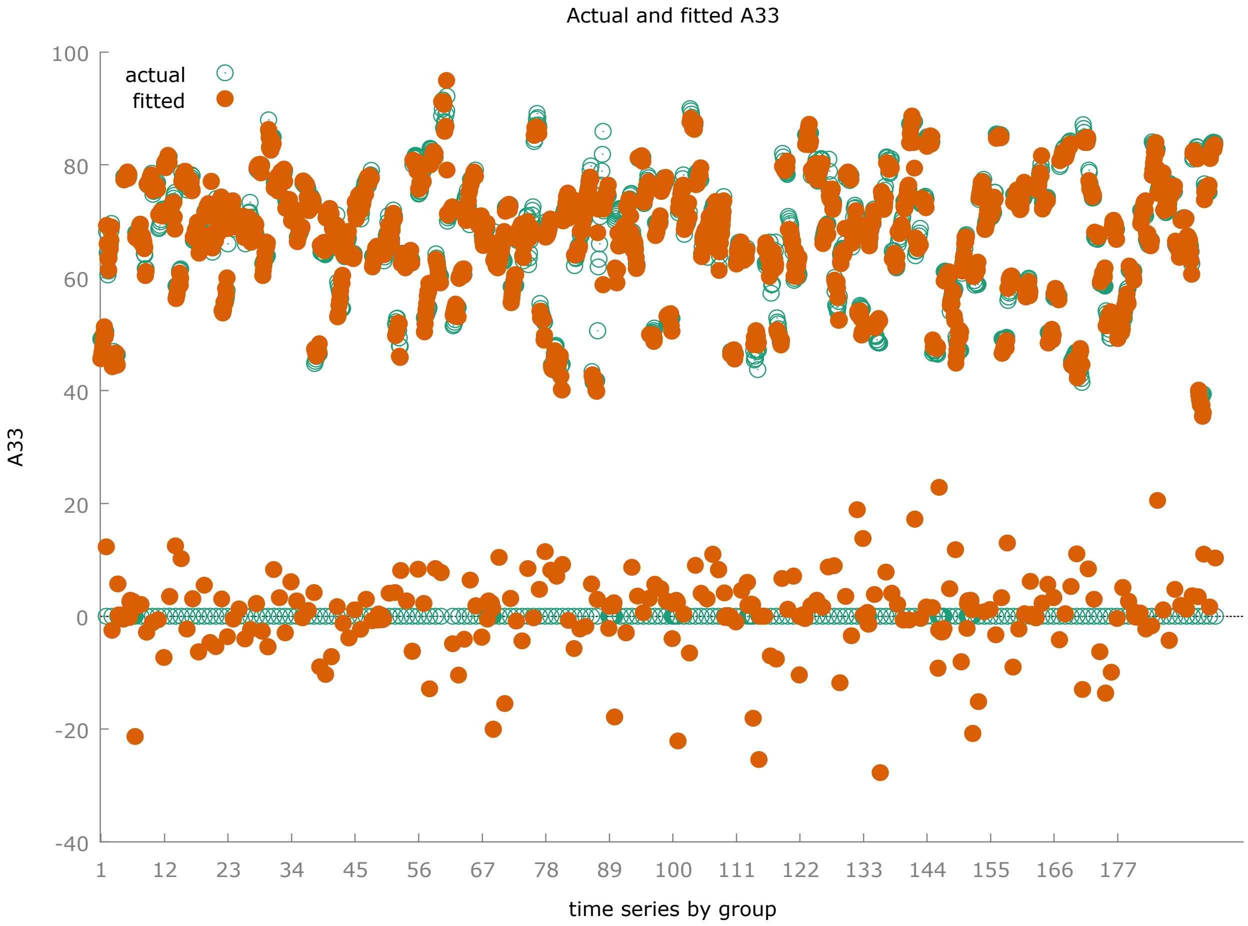
with p-value = 8,02221e-011



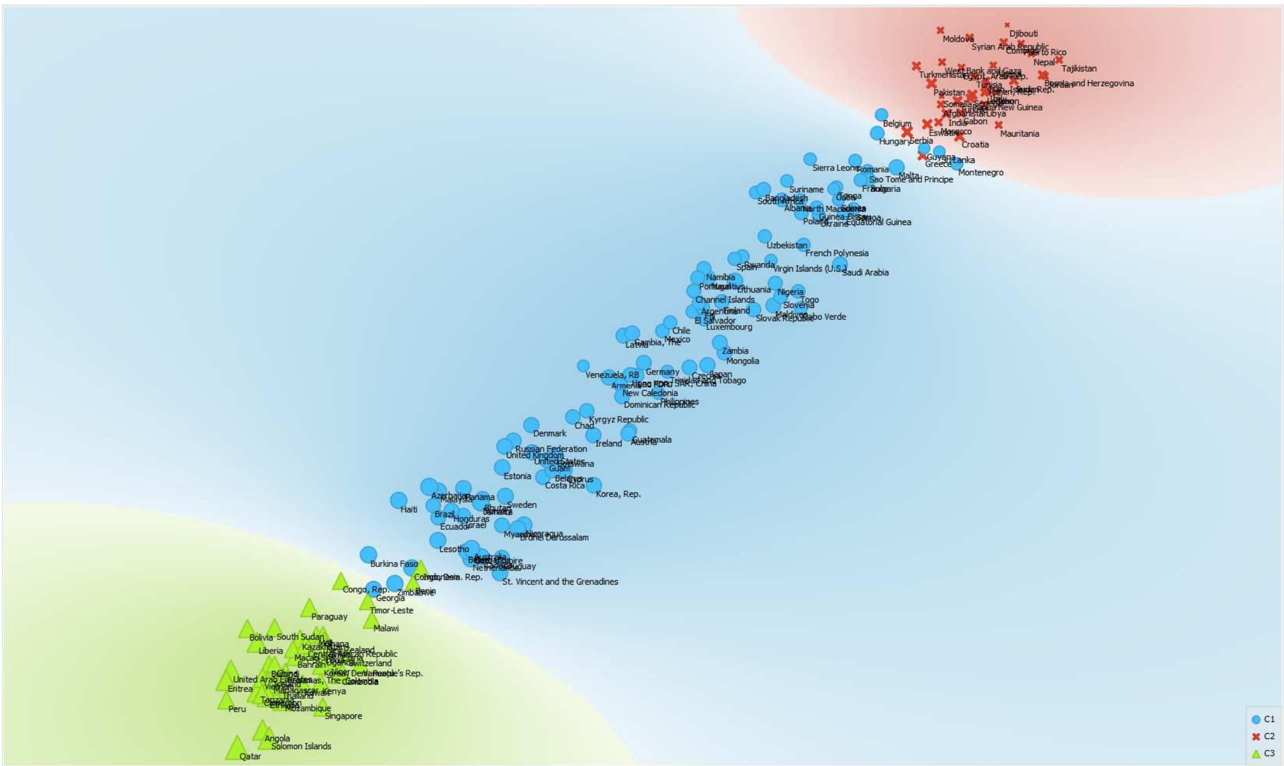
Fixed-effects, using 1930 observations					
Included 193 cross-sectional units					
Time-series length = 10					
Dependent variable: A33					
	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	3,61957	0,478375	7,566	<0,0001	***
A5	−0,0145975	0,00646848	−2,257	0,0242	**
A11	0,133879	0,0376956	3,552	0,0004	***
A34	0,296225	0,0109720	27,00	<0,0001	***
A46	0,0402140	0,00675652	5,952	<0,0001	***
A54	0,623847	0,0106825	58,40	<0,0001	***
A55	−0,0892834	0,0228191	−3,913	<0,0001	***
A65	−0,344686	0,0559300	−6,163	<0,0001	***
Mean dependent var	56,37973	S.D. dependent var	27,26489		

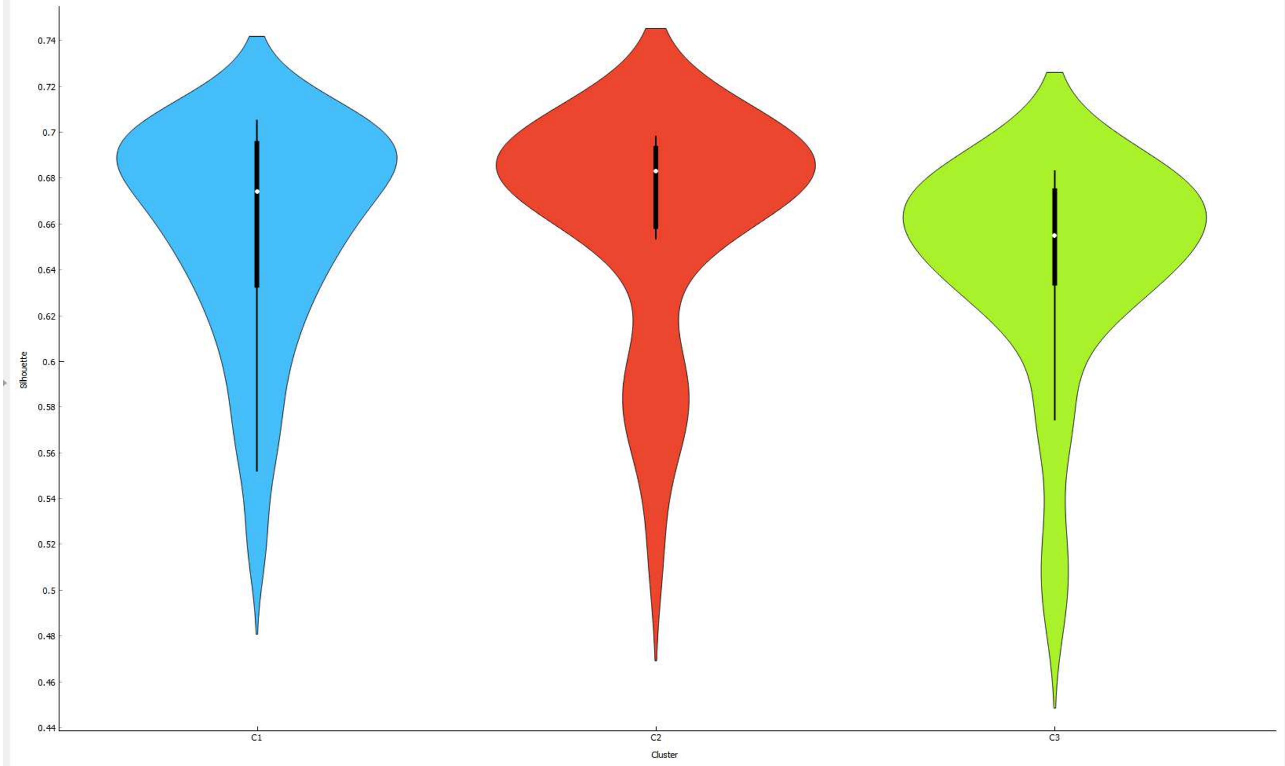
Sum squared resid	22625,14	S.E. of regression	3,616368
LSDV R-squared	0,984222	Within R-squared	0,969713
LSDV F(199, 1730)	542,2936	P-value(F)	0,000000
Log-likelihood	−5113,939	Akaike criterion	10627,88
Schwarz criterion	11740,93	Hannan-Quinn	11037,31
rho	0,286115	Durbin-Watson	1,064062

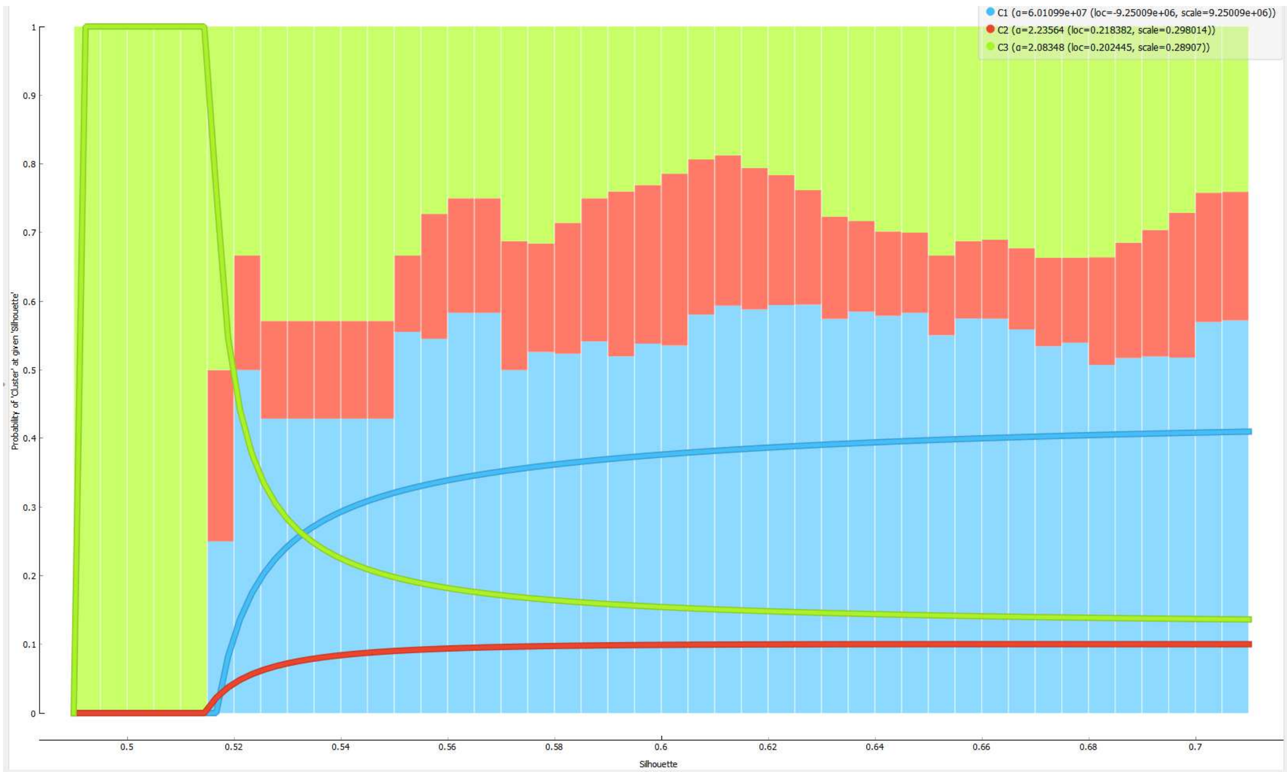
Joint test on named regressors -
Test statistic: F(7, 1730) = 7913,01
with p-value = P(F(7, 1730) > 7913,01) = 0
Test for differing group intercepts -
Null hypothesis: The groups have a common intercept
Test statistic: F(192, 1730) = 56,6111
with p-value = P(F(192, 1730) > 56,6111) = 0



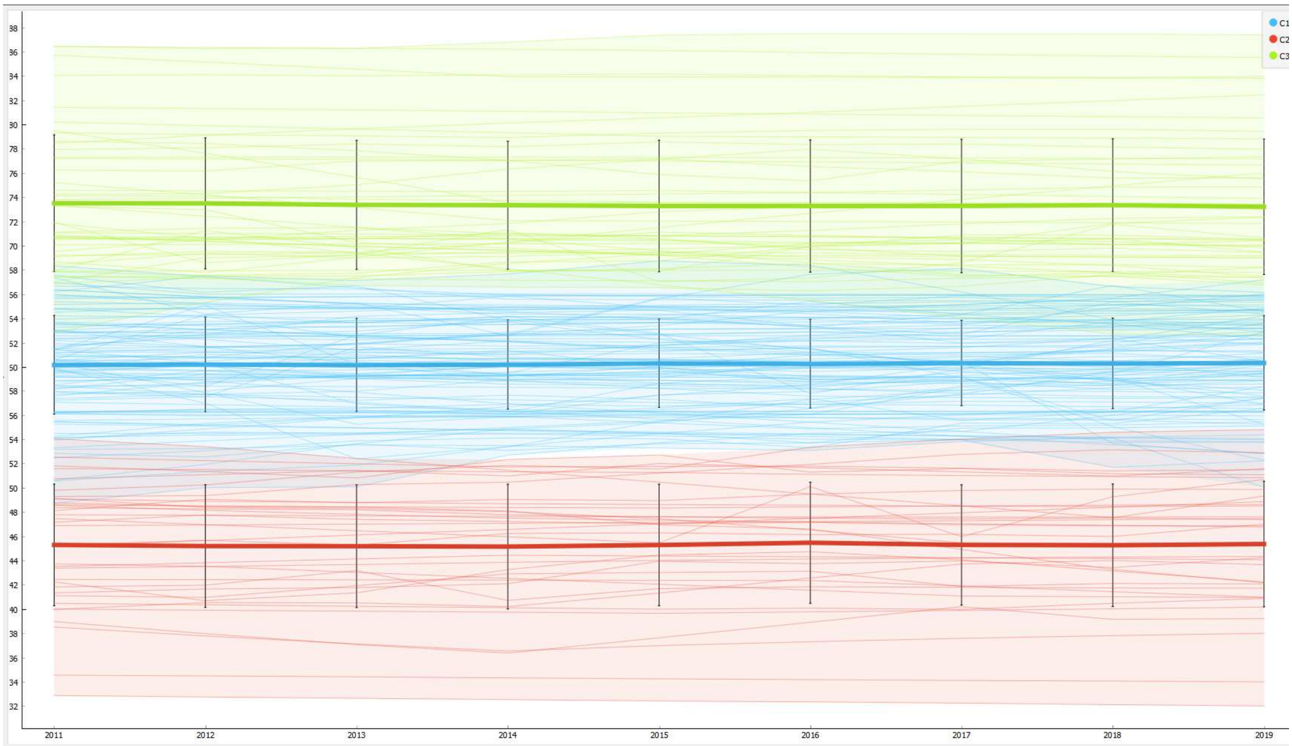

0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
C1 (107)
Slovak Republic
Channel Islands
Chile
El Salvador
Czechia
Argentina
Finland
Luxembourg
Portugal
Gambia, The
Lao PDR
Zambia
Latvia
Japan
Mexico
Mongolia
Trinidad and Tobago
Germany
Mauritius
Hong Kong SAR, China
Namibia
Austria
Lithuania
Spain
Philippines
Cabo Verde
Fiji
Togo
Armenia
Guatemala
Denmark
Ireland
Chad
Botswana
Slovenia
Estonia
Dominican Republic
Venezuela, RB
United States
Rwanda
Guam
Kyrgyz Republic
Russian Federation
Nigeria
United Kingdom
Uzbekistan
South Africa
Korea, Rep.
Bangladesh
Cyprus
New Caledonia
Maldives
Costa Rica
Poland
Virgin Islands (U.S.)
Belarus
French Polynesia
Albania
Ukraine
Brazil
Guinea-Bissau
Honduras
Israel
Azerbaijan
Saudi Arabia
Belize
Sierra Leone
Cuba
Panama
France
Sweden
Equatorial Guinea
Bhutan
Myanmar
Malaysia
Jamaica
Guinea
Netherlands
North Macedonia
Norway
Suriname
Malta
Cote d'Ivoire
Brunei Darussalam
Uruguay
Nicaragua
St. Vincent and the Grenadin
Barbados
Ecuador
Australia
Romania
Lesotho
Tonga
Sao Tome and Principe
Samoa
Bulgaria
Guyana
Hungary
Zimbabwe
Haiti
Canada
Congo, Dem. Rep.
Belgium
Burkina Faso
Georgia
Montenegro
Sri Lanka
C2 (35)
Syrian Arab Republic
Comoros
West Bank and Gaza
Tajikistan
Iraq
Moldova
Algeria
Mauritania
Bosnia and Herzegovina
Iran, Islamic Rep.
Puerto Rico
Nepal
Afghanistan
Tunisia
Lebanon
Turkmenistan
Libya
Morocco
Egypt, Arab Rep.
Papua New Guinea
Jordan
Yemen, Rep.
Senegal
Gabon
Sudan
Somalia
Italy
Djibouti
India
Eswatini
Pakistan
Turkiye
Croatia
Greece
Serbia
C3 (45)
Korea, Dem. People's Rep.
Kenya
Vietnam
Bahamas, The
Liberia
Angola
Iceland
Peru
Eritrea
Niger
Cameroon
Burundi
Mozambique
Bahrain
Kuwait
United Arab Emirates
Ethiopia
South Sudan
Central African Republic
Kazakhstan
Macao SAR, China
Uganda
Solomon Islands
Cambodia
Tanzania
Singapore
Mali
Madagascar
Ghana
Qatar
Paraguay
Malawi
Vanuatu
St. Lucia
Bolivia
Thailand
China
New Zealand
Colombia
Switzerland
Congo, Rep.
Oman
Timor-Leste
Indonesia
Benin
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
C1
C2
C3

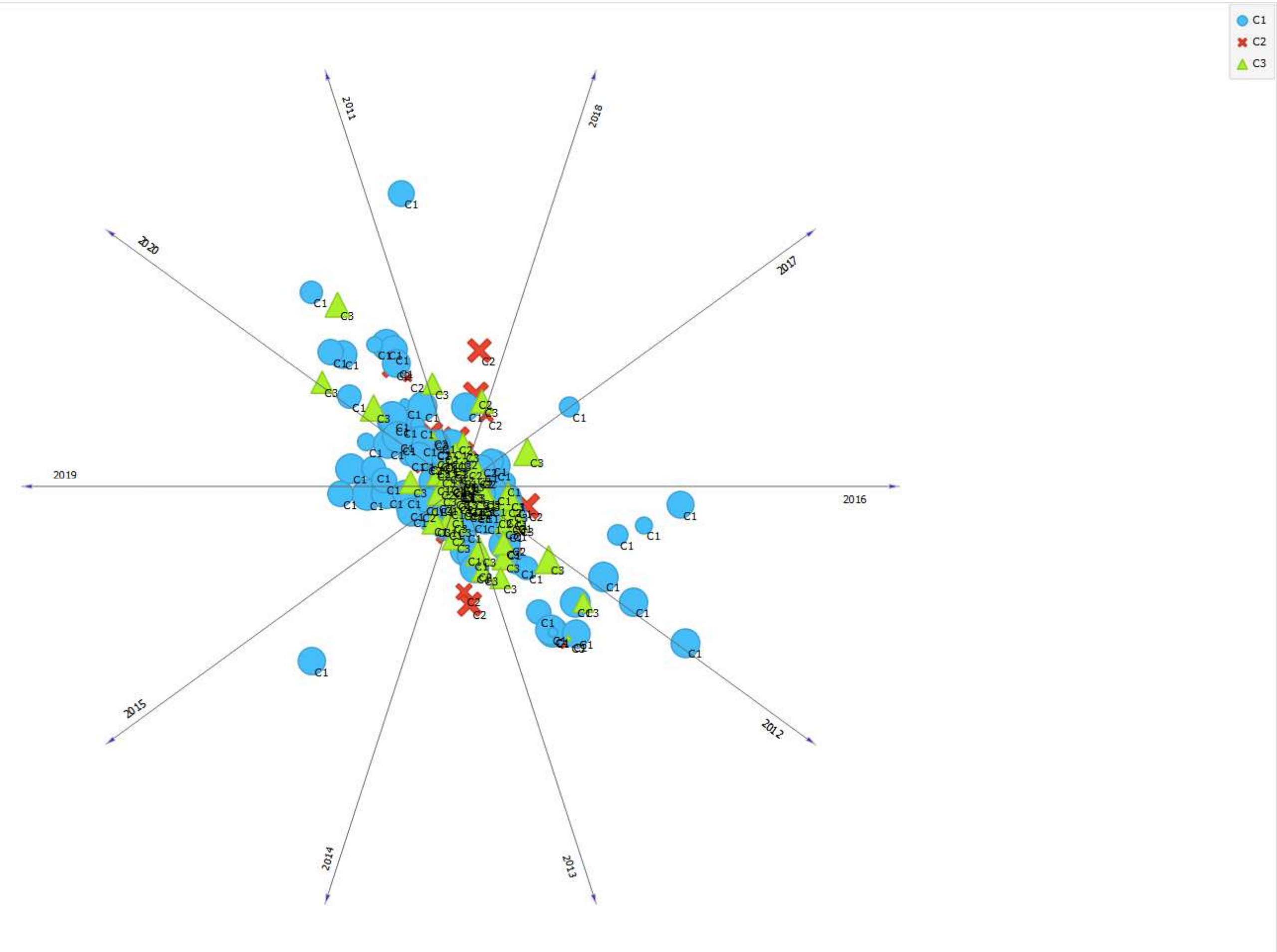

C1
C2
C3

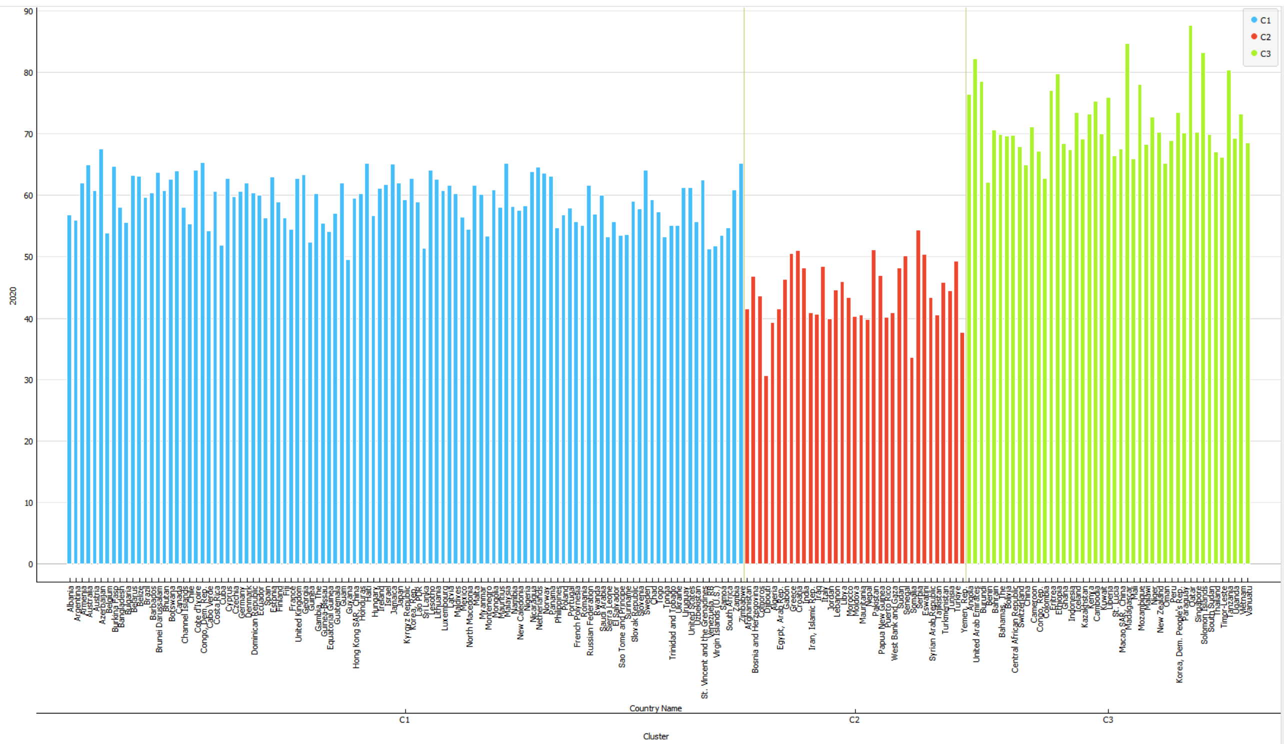

Silhouette
0.74
0.72
0.7
0.68
0.66
0.64
0.62
0.6
0.58
0.56
0.54
0.52
0.5
0.48
0.46
0.44
C1
C2
C3
Cluster

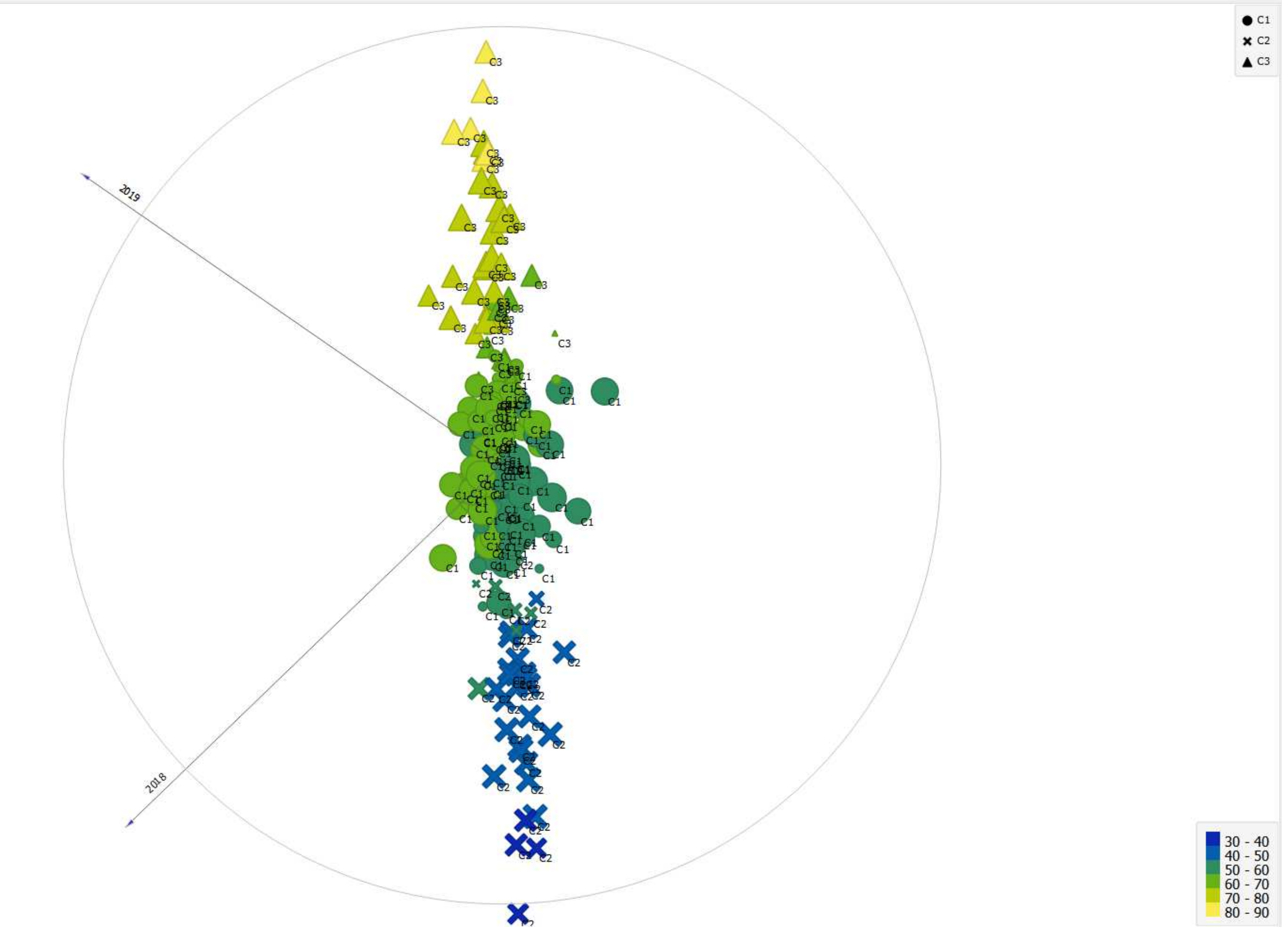

C1 (α=6.01099e+07 (loc=-9.25009e+06, scale=9.25009e+06))
C2 (α=2.23564 (loc=0.218382, scale=0.298014))
C3 (α=2.08348 (loc=0.202445, scale=0.28907))
Probability of 'Cluster' at given 'Silhouette'
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.5
0.52
0.54
0.56
0.58
0.6
0.62
0.64
0.66
0.68
0.7
Silhouette

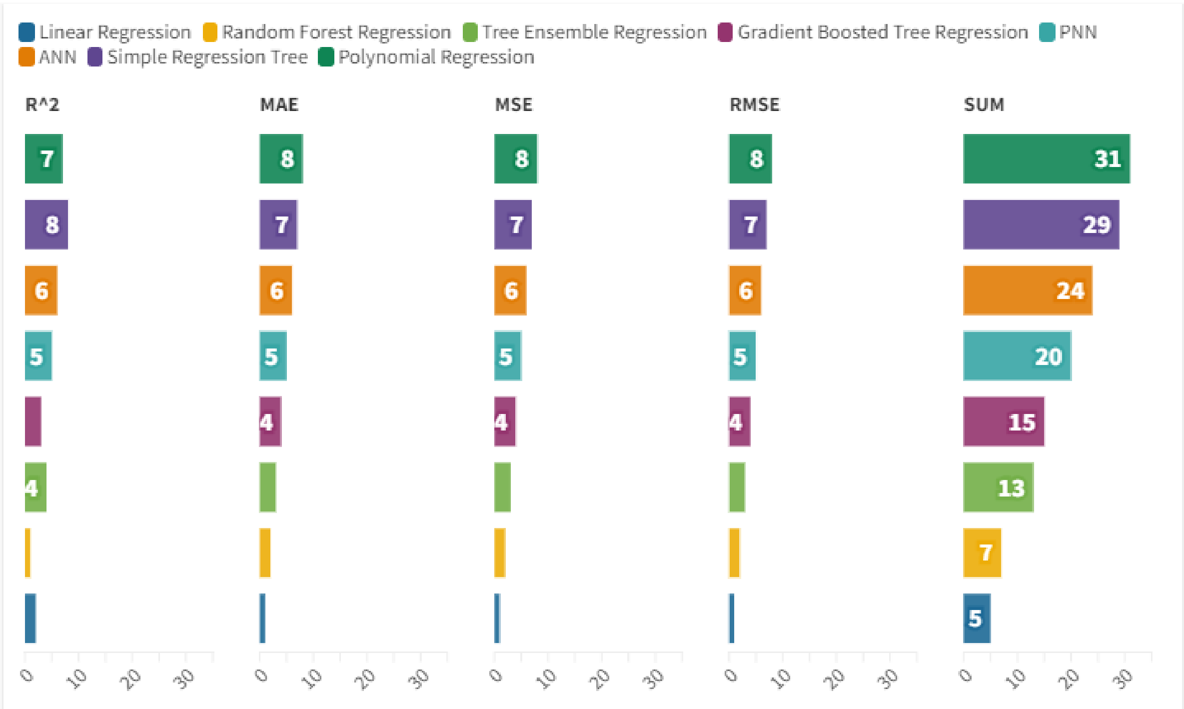

Silhouette
Cluster
C1
C2
C3
Namibia
Lithuania
Slovak Republic
Portugal
Mexico
Armenia
Luxembourg
Rwanda
Hong Kong SAR, China
Mongolia
Japan
Togo
Cabo Verde
Lao PDR
Fiji
Uzbekistan
Estonia
Cyprus
Belarus
Korea, Rep.
Virgin Islands (U.S.)
Nigeria
Maldives
France
Azerbaijan
Brazil
Costa Rica
Panama
Israel
French Polynesia
Sweden
Malaysia
Honduras
Bhutan
Equatorial Guinea
Malta
Suriname
North Macedonia
Belize
Jamaica
Guinea
Myanmar
Brunei Darussalam
Nicaragua
Netherlands
Cote d'Ivoire
Uruguay
Ecuador
Romania
Lesotho
Barbados
Australia
St. Vincent and the Grenadines
Tonga
Sao Tome and Principe
Guyana
Haiti
Bulgaria
Hungary
Samoa
Belgium
Zimbabwe
Canada
Burkina Faso
Congo, Dem. Rep.
Georgia
Sri Lanka
Montenegro
Moldova
Syrian Arab Republic
Comoros
West Bank and Gaza
Algeria
Tajikistan
Puerto Rico
Nepal
Bosnia and Herzegovina
Turkmenistan
Egypt, Arab Rep.
Afghanistan
Lebanon
Jordan
Sudan
Morocco
Gabon
Djibouti
Italy
Somalia
India
Pakistan
Eswatini
Turkiye
Croatia
Greece
Serbia
Liberia
Korea, Dem. Peo
Niger
Kenya
Eritrea
Peru
Bahrain
Mozambique
Cameroon
Central African Republic
Kuwait
Macao SAR, China
Uganda
Madagascar
Cambodia
Ghana
Malawi
Bolivia
Solomon Islands
Singapore
St. Lucia
Vanuatu
Qatar
New Zealand
China
Thailand
Colombia
Congo, Rep.
Oman
Switzerland
Timor-Leste
Indonesia
Benin


C1
C2
C3
2011
2012
2013
2014
2015
2016
2017
2018
2019


C1
C2
C3
2011
2018
2020
2017
2019
2016
2015
2012
2014
2013


C1
C2
C3
90
80
70
60
50
40
30
20
10
0
2020
Country Name
Cluster


C1
C2
C3
2019
2018
30 - 40
40 - 50
50 - 60
60 - 70
70 - 80
80 - 90


Linear Regression
Random Forest Regression
Tree Ensemble Regression
Gradient Boosted Tree Regression
PNN
ANN
Simple Regression Tree
Polynomial Regression
R^2
MAE
MSE
RMSE
SUM
7 8 6 5 4
8 7 6 5 4
8 7 6 5 4
8 7 6 5 4
31 29 24 20 15 13 7 5
0 10 20 30