The impact of the population's age composition on technological progress

Homonenko, Vladyslava and Suprun, Ivan and Platonovska, Vladyslava

Kyiv School of Economics

1 April 2024

[1] The impact of the population's age composition on technological progress

Homonenko Vladyslava, Platonovska Vladyslava, Suprun Ivan

Abstract

This research paper explores the determinants of countries' innovation levels as measured by the Global Innovation Index, focusing on the essential factors that strengthen a country's innovative capabilities. Through a comprehensive cross-country regression analysis, the findings highlight the role of GDP per capita, Median age and the Share of the population aged 25-49 as significant factors of innovation. Contrary to our initial hypotheses, the analysis demonstrates that other variables such as the democracy level, birth rate, net migration, and life expectancy, initially provided in our model, do not significantly influence the innovation process. This indicates that economic prosperity, a youthful age profile, and a significant part of the population within their most productive years are crucial in enhancing a nation's innovation. The results are particularly significant for Ukraine, emphasizing the need to enhance technological progress in its post-war recovery efforts. This study confirms the importance of a relatively young population and stable economic health in fostering technological progress, providing guidance for policymakers aiming to enhance innovation strategies.

Keywords: Global Innovation Index, Technological progress, Age distribution, Economic health.

Introduction

To start with, our aim is to investigate how the age distribution across the particular countries may influence technological progress. Our team is interested in either confirming or rejecting the theory of the fact that countries with the bigger part of the population with younger people, rather than older ones, have more rapid technological progress and develop quicker in terms of new innovations. We will choose a specific number of countries, along with the information of age demographics and their technological improvements.. Furthermore, we will take several characteristics, that are our future independent variables to make conclusions about our dependent variable, which is technological progress across the country. This study will be a comparable one, as we are to compare different countries and different age ranges as well. We hope that our analysis will provide useful information for policymakers, educators or business makers on how to apply demographic trends to provide more technological advancement and enhance the innovation of new products.

---

[1] This study was part of the Regression Analysis course at Kyiv School of Economics, conducted under the curriculum designed for second-year students.

Diving deeper into the relationship between age demographics and technological progress unveils a complex web of factors that influence a nation's trajectory in the digital age. This topic is important for several reasons. First of all, it emphasizes how important human capital is to creating innovation ecosystems. A younger population can greatly speed up the adoption of new technologies, from blockchain to artificial intelligence, influencing a nation's economic future and standing in the international community. Younger people are also frequently more adaptive and digital native. Moreover, this inquiry into demographic influences extends beyond mere economics into the realm of social policy and education. It raises key issues about how educational institutions should be set up to ensure that the next generation is not simply a consumer of technology, but also an innovator and creator in a society where technology is dominant. Additionally, the aging population's impact on technology demands attention. The focus on adaptive technologies, healthcare innovations of an older generation presents an ideal foundation for technological growth, making it a crucial topic of interest for politicians and entrepreneurs alike. This demographic instability is something that many advanced countries are facing.

Finally, understanding these dynamics provides a framework for analyzing the biggest societal transformations caused by technological innovation, such as changes to the structure of communities, the nature of work and labor markets. As a result, the relationship between age demographics and technological advancement is not only a matter of research curiosity, but also a puzzle that must be solved.

The results of our study may also provide important implications for the case of Ukraine. The country is known for its demographic problems that have been worsened by the current conflict with Russia. If our analysis supports the conventional theory that younger people tend to generate more technological progress, we might be able to propose some policies to improve demographics in our country. Some additional parameters that we will choose might be significant as well, so we can pay attention to health care or something like that to improve in order to achieve a higher rate of innovations.

Data and Methods Description

Data sources

We have chosen several datasets with precise information from different sources for our deep research about the influence of age distribution across the country on its Innovation Index. Most of them was obtained from the World Bank, which is the one of the premier and most trustworthy sources of data, containing the most credible and relevant information for our research. However, we have also included datasets from other sources, such as the United Nation and Central Intelligence Agency website. Before deployment, the data went through thorough cleansing and quality assurance processes to ensure its accuracy to prepare it for our analysis.

Dependent variable

The dependent variable for our research question is Global Innovation Index, which is the most relevant measure to choose for investigation of innovation progress across the countries, expressed as a score on a scale ranging from 0 to 100. GII is calculated by taking the mean value of 2 sub-indices : Innovation Input

Index and Innovation Output Index, which are composed of five most vital for innovation indicators: institutions, human capital and research, infrastructure, knowledge and technology outputs, and creative outputs. In this way, GII measures the intensity of innovation processes in a country by evaluating how these key areas contribute to creating an environment conducive to innovation. By assessing both the inputs that facilitate innovation, such as the quality of institutions, the level of human capital and research, and the state of infrastructure, and finally, the outputs that result from innovation efforts as well, estimated by advancements in knowledge and technology, as well as creative products, it provides a clear picture of a nation's innovation ecosystem.

Independent variables

The chosen independent variables for our investigation, along with their respective metrics, are as follows:

1. **GDP per capita (current US$)** : The metric, which is determined by dividing a country's GDP by its population size, serves as an indicator of overall development and productivity of the nation.

2. **Life expectancy (years) :** This variable estimates the average number of years a person is expected to live based on the current mortality rates and health conditions in a country. It shows a clear picture of public health, welfare, and social and economic conditions.

3. **Birth rate (births/1,000 population) :** Birth rate provides the average annual number of births during a year per 1,000 persons in the population at midyear.

4. **Age dependency ratio (% of working-age population) :** It is a measure expressed as the percentage of the labor force, which is aged 15-64 and indicates the proportion of people, who are younger than 15 and older than 64 in a country.

5. **Country democracy :** shows whether the country is democratic or not.

6. **Median age (years)** : Statistical measure that describes the middle point of the age distribution of the population of the country.

7. **Share of population aged from 15-24 and 24-49 years (in %)** : This measure shows the percentage of the country's population among this certain range. We chose this specific range to be included to explore whether the percentage of younger population influency on the technological development in the country.

8. **Net migration (amount of people) :** The difference between the number of immigrants and the number of emigrants divided by the population.

Our independent variables were selected to be in our regression analysis due to their theoretical significance for innovation processes, so that we will investigate further whether they are really statistically significant for our investigation.

Regression model

Our multiple linear regression model, which will help us to identify the dependency of age distribution across the country on its technological development and innovations, represented as :

> *Global Innovation Index = f (GDP per capita, Life expectancy, Birth rate, Age dependency ratio, Democracy, Median Age, Share of people aged 15-24, Share of people aged 24-49, Net migration)*

Embellishment of independent variables

We selected our independent variables based on the assumptions that they are vital indicators of a country's citizen welfare, living standards, and capacity for innovation and development. These factors are pivotal in driving national progress and ensuring the overall stability and growth of the country. We are here to provide our assumptions of how those factors can influence the country's process of qualitative advancement in terms of innovation intensity.

1. **GDP per capita** : It is an absolutely crucial indicator to understand the economic health of the country and its overall standard of living. It may also indicate the amount of funds, that are available for R&D investment. Therefore, the higher GDP per capita often correlates with better technological infrastructure and a greater capacity for technological progress.

2. **Life expectancy** : It serves as a proxy for the health among the population. Healthy citizens are more productive and ready for the new ideas to imply, which may contribute to the country's technological advancement and overall GII.

3. **Birth rate** : The rate may show us the part of the younger population, as the birth rate influences the age structure of the population.

4. **Age dependency ratio** : The ratio will provide us the information about percentage of those who are in the labor force and those who are dependents, which is useful in our model to understand the percentage of people who have the ability to contribute in country's technological advancement.

5. **Country democracy** : Democratic countries tend to support a freedom of expression, provide policies, which include support for R&D, intellectual property rights protection, and the promotion of entrepreneurship. Having those advantages ease and enhance the country's rate of innovation.

6. **Median age** : The demographic maturity of the country is our the most important factor to examine. We will analyze whether lower median age, supposing a larger part of young population, affects the speed of country's technological growth.

7. **Share of population aged from 15-24 and 24-49 years :** This percentage is interesting for us to investigate whether a higher proportion of the young population is contributing to innovation and technological development.

8. **Net migration** : Indicates the attractiveness of a country in terms of opportunities, which can be a proxy for its innovative capacity and technological advancement.

Prior expectations

The expectations about how each predictor will impact the GII, based on our initial theories, can be articulated in the following manner:

Hypothesis 1: Democratic countries have higher scores on the Global Innovation Index than non-democratic countries.

Hypothesis 2: An increase in GDP per capita is positively related to the Global Innovation Index.

Hypothesis 3: A higher average life expectancy is associated with higher scores on the Global Innovation Index.

Hypothesis 4: A higher birth rate is positively associated with the Global Innovation Index, as younger population stimulates greater innovation.

Hypothesis 5: Age Dependency Ratio is negatively related to the Global Innovation Index, as a higher ratio of dependents to the working-age population reduces available resources and labor.

Hypothesis 6: Lower median age indicates a more dynamic and rapidly developing population, influencing the index positively.

Hypothesis 7: A higher proportion of the population aged between 15-24 and 24-49 is to have a positive impact on the Global Innovation Index, reflecting a demographic more inclined towards innovation and technological advancement.

Hypothesis 8: A higher rate of net migration negatively influences the Global Innovation Index, as people often migrate towards countries offering greater opportunities and a more developed technological sector, leaving their home-country, which has less rate of innovations.

Model estimation methodology

The model will be analyzed by using the Ordinary Least Squares regression method. Ordinary Least Squares regression analysis aims to create a final model that accurately represents the relationship between the dependent and its independent variables. This is achieved by reducing the total of squared deviations between the actual and predicted values of the dependent variable, which are in the fitted regression model. By estimating the model by OLS method, the residuals, which are those differences, will be minimized. Moreover, the Ordinary Least Squares regression method will provides us with the F-test and adjusted R2 to estimate the goodness-of-fit of our model to ensure we include all the needed predictors for the model's strong efficiency.

Model equation

Our suggested linear multiple regression equation is as it follows :

Global Innovation Index = β0 + β1 · GDP_ per_capita + β2 · democratic + +β3 · median_age + β4 · birth_rate + β5 · net_migration +β6 · life_expectancy + β7 · population_aged_15-24 + β8· population_aged_24-49 + β9 · age_dependency_ratio + ε

Where:

- "β0" is the intercept.
- "βi" are the coefficients for our predictor variables, such as : GDP per capita,democracy, median age, life expectancy, share of population aged 15-24, share of population aged 24-49, net migration, age dependency ratio and the birth rate.
- "ε" is the error term.

Further, we will obtain the final model by exploring all the variables and excluding those that are not statistically significant, if needed.

Empirical Analysis and Results

Visualisation of the relationships between variables

Firstly, we have transformed our data for our variables to be numeric and recognisable by R to proceed with the analysis correctly further. After completing the required transformations, we are moving to the exploration of the relationships between all the variables visually.

The scatterplots provided below (see Figure 1) shows the following results :

1. There is a positive linear relationship between the Global Innovation Index and the GDP per capita of a country, suggesting that a higher GDP per capita is associated with the higher GII among the country.
2. There is a positive linear relationship between the Median Age of a country's population and its GII, indicating that countries with higher average age tend to have higher GII scores.
3. Life expectancy also appears to correlate positively with GII, indicating that improved health outcomes may potentially increase GII.
4. Global Innovation Index, Democracy, Net Migration and Share of population aged 15-24, as well as group of those who are between 25 and 49, do not have a strong relationship visually, as the points are scattered randomly without any potential explanation.
5. Birth rate seems to have a negative relationship with GII, implying that countries with higher birth rates may have lower GII scores.

Also, on the scatter plot, there seems to be no relationship between independent variables, so we can expect the lack of the problem of collinearity in our analysis.

Calculating the correlation's coefficients

After we created a visualization of the relationship between all the variables, we have to explore the relationships with the analytical method, so we need to proceed with calculating the correlation between all the variables (see Table 1).

When exploring the correlation coefficients (see Table 2), statistics show that there is a very strong positive relationship between the dependent variable, GII and GDP per capita (0.8), Life Expectancy (0.8). Also, we can see a strong negative relationship between GII and Share of population aged 15-24 (-0.8) and between GII and Birth rate (-0.8). However, there is a relatively weak correlation between our dependent variable and Democracy, as well as with Net Migration, so we can expect to exclude those variables from our model, while proceeding with choosing the final model (see Table 3). However, as it was suggested visually, there would be no problem of the collinearity in our model, as any of the correlation's coefficients for the pairs of independent variables do not exceed or equal to 0.9.

Choosing the final model

We start with the full model by including all the selected predictors. The initial model is as follows :

```
Model_full <- lm(`Global Innovation Index`~ Democratic + `GDP Per capita 2021` + `Median Age` + `Birth rate` + `Net migration` + `Life Expectancy` + `15-24` + `25-49`)
```

However, we use the backward selection in order to determine our final regression model without any not statistically significant predictors, which will be utilized to evaluate the impact on the Global Innovation Index. We will exclude all the variables that exceed the chosen significance level of 5% one-by-one.

After completing the process of the backward elimination (see Table 4), it appears that there are only 4 predictors being statistically significant in terms of their effect on the GII at the 5% significance level. The final model we got is as:

```
model_reduced <- lm(`Global Innovation Index`~ `GDP Per capita` + `Median Age`  + `Life Expectancy` + `25-49`)
```

However, for better analysis we also need to check for potential interactions within our model, as some variables may have different effects on GII, when interacting with another predictor, so that the result can be changed significantly.

After checking for the statistically significant interactions within our independent variables, we have the new model, which includes the interaction term between GDP Per capita 2021 and Share of population aged 25-49 :

```
model_with_interaction <- lm(`Global Innovation Index` ~ `Median Age` +`GDP Per capita 2021` * `25-49` + `Life Expectancy`)
```

The next step is to compare the models that we have already received.

```
model_full <- lm(`Global Innovation Index`~ Democratic + `GDP Per capita 2021` + `Median Age` + `Birth rate` + `Net migration` + `Life Expectency` + `Group 15/24` + `Group 25/49`)

model_final <- lm(`Global Innovation Index`~`GDP Per capita 2021` * `Group 25/49` + `Median Age` +`Life Expectancy`)
```

The ANOVA test (see Table 7), which compares two models showed that there is no significant difference in its ability to explain the variance in the dependent variable. With a high p-value of 0.9139, the additional variables included in model_full do not significantly enhance the model's explanatory power compared to the model_final. Additionally, an examination for multicollinearity using the vif() function (see Table 5) yielded no values exceeding the threshold of 5, confirming the fact that multicollinearity is not a concern in our final model. Consequently, the streamlined model is deemed to be the most appropriate and effective for the analysis.

Before interpreting the results, it's essential to check the remaining linear regression assumptions. Previously, the data transformation was employed to ensure compliance with the linearity assumption. Constant variance in the residuals was visually verified and further supported by the Breusch-Pagan test, which resulted in a p-value above the significance threshold of 5%. Nonetheless, the Shapiro-Wilk test indicated that the residuals did not conform to a normal distribution, thus not meeting this particular assumption. To correct for this violation, we apply the log-transformation to our dependent variables. Through these adjustments, we not only attain normality in the residual distribution, as indicated by a p-value greater than 0.05, but we also see a rise in the R-squared value, increasing from 0.82 to 0.84.

After applying the log transformation to GII, the final model (see Table 6) can be presented as follows:

```
model <- lm(log(`Global Innovation Index`) ~`GDP Per capita 2021`+ `Median Age` + `Group 25/49` + `GDP Per capita 2021`:`Group 25/49` )
```

Based on the final model summary, the fitted regression equation is as follows:

log (Global Innovation Index) = 2.281 + 0.001* GDP per capita 2021 + 0.019* Median Age + 0.012 * Group 25/49 - 0.001 * GDP Per capita * Group 25/49

According to the Adjusted R-Squared coefficient, our model explains approximately 84.4% of all cases, nevertheless, we must point out that there are some unusual observations we kept in our model. Those are: 13 – Bahrein, 36 – China, 80 – Ireland, 81 – Israel, 100 – Luxembourg, 129 – Norway, 141 – Qatar, 168 – Sweden, 185 – United Arab Emirates. We decided to keep them as those all seem very important to keep for better understanding of the real relationship. In our opinion, those countries are so well-known and developed in terms of its GII, that neglecting their impact on the model is not scientifically right.
All the coefficients in our model are positive, except for that on the interaction term, which is very small if we round numbers to two decimal places. All the other coefficients are quite small but still significant and have a positive influence on our dependent variable.

Before moving to the analysis, we had initial assumptions about the outcome. The hypotheses we had are as follows :

Hypothesis 1: An increase in GDP per capita is positively related to the Global Innovation Index.

Hypothesis 2: Lower median age indicates a more dynamic and rapidly developing population, influencing the index positively.

Hypothesis 3: A higher proportion of the population aged between 24 and 49 is to have a positive impact on the Global Innovation Index, reflecting a demographic more inclined towards innovation and technological advancement.

Compared to the theoretical expectations and logical predictions, the coefficients' signs align well with almost all anticipated outcomes, even though they are small. This aligns with expectations when analyzing the varying levels of innovation across countries. Unexpectedly though, an increase in the Median Age leads to a positive effect on the Global Innovation Index (GII), suggesting that this effect may remain positive up to a certain point, after which its impact might diminish.

Additionally, an increase in the share of the population aged 25-49 positively influences the GII, although the extent of this impact varies with the country's level of economic development. As countries become wealthier, the positive effect diminishes. To illustrate this, we conducted calculations for three groups of countries based on their economic development levels, using the partial derivative with respect to the share of the population aged 25-49:

Dy/D(25-49) = 0.014178 - 0.004040 * GDP Per Capita.

For countries at a relatively low development level with a GDP per capita of about $10,000, the impact of the share of the population aged 25-49 on GII is calculated as :
0.014178 - 0.004040 * 10,000 = -40.4 (approximately).

Countries at an average level of economic development, with a GDP per capita of about $50,000, see an impact of -202 (approximately).

For very wealthy countries with an average GDP per capita of about $150,000, the impact is calculated as -606 (approximately).

Those calculations reveal that the significance of the population group aged 25-49 on the GII declines as countries become wealthier. This suggests that this demographic is more crucial for the innovation processes in less economically developed countries and less significant in more developed economies. Therefore, it can be inferred that wealthier nations require a younger population to foster a more vibrant innovation ecosystem.

Conclusions

To conclude, our regression analysis has explored the dynamics between a country's age composition and its development in terms of technological innovation, specifically within looking at the Global Innovation Index among several chosen countries. The study highlighted the role of GDP per capita, Median age and Share of population aged 25-49 years as critical determinants of a country's innovative progress. Oppositely to our preconceived assumptions, the statistical analysis did not confirm any evidence to suggest an impact of other factors, such as : Democracy of a nation, Birth rate, Net migration and Life expectancy, that were initially considered should be in our regression model. Therefore, those insights that we got in our final model highlight the importance of economic health, young population among the country, and overall country's well-being in technological progress.

Based on these findings, several policy recommendations are to be presented. The most important result we have obtained is that it is necessary to enhance policies that are aimed at boosting economic growth and enhancing GDP per capita, as those factors are the most crucial to have the constant technological development in the country. This process would include strategies of maintaining economic stability, which may be done through a reasonable mix of Monetary and Fiscal policies. Those might include measures such as adjusting the money supply, tax rates, or managing government budgets. Additionally, such efforts as attracting investments, boosting the overall productivity, and improving the education system for future generations are essential for improving the country's GDP per capita. Moreover, the results showed that countries with higher GDP per capita are in need of having a higher share of young population among the whole nation's population, as it would increase productivity and innovation progress more. Consequently, to enhance the innovation progress in wealthy counties, policymakers may consider investing in lifelong learning, ensuring that the labor force remains innovative and adaptable. Fostering Research and Development is also crucial, as it is the most important factor for the country to grasp new opportunities for potential technological growth. Additionally, creating internship programs where experienced workers guide younger employees could enhance knowledge and foster innovation.

For Ukraine, particularly, applying those insights is essential to understand the needed strategy for the post-war development. Given that Ukraine's GDP per capita is relatively low, it's crucial for the government to implement the policies written previously to enhance economic well-being. As a relatively low-developed country, compared to others, Ukraine has a bigger impact of the younger population on innovation. However, it's equally important to focus on improving the overall economic well-being and stimulate GDP per capita to have more opportunities for future development.

Moreover, we need to continue and enhance that dynamic energy of the Ukraine's young population by investing in qualitative education, particularly in STEM fields, and provide useful and available opportunities, such as free courses to apply this knowledge in well-paid jobs. This dual approach of boosting the economy while giving the possibilities for youth to study life-needed subjects can serve as a powerful scenario of future progress for Ukraine.

Further research can be aimed to explore additional factors that may influence GII, for example, political stability, digital infrastructure or intellectual property rights. Moreover, conducting analyses of some particular countries separately may bring insightful results about the predictors and its effects on the Global Innovation Index. We plan to conduct an analysis approximately a year later after Ukraine's to observe the expected increase in the GII. Through regression analysis, we will identify which factors have contributed to this improvement, providing insightful results about the drivers of innovation post-conflict.

References

1. CIA The World Factbook (2022). Archives
https://www.cia.gov/the-world-factbook/about/archives/2022/

2. CIA The World Factbook (2023). Median Age
https://www.cia.gov/the-world-factbook/field/median-age/country-comparison/

3. UNECE (2021). Labor force by age and sex
https://w3.unece.org/PXWeb2015/pxweb/en/STAT/STAT__30-GE__03-WorkAndeconomy/001_en_GEWELabourActivity_r.px/

4. United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022, Online Edition.

5. The World Bank (2021). Life expectancy at birth, total (years)
https://data.worldbank.org/indicator/SP.DYN.LE00.IN

6. The World Bank (2022). GDP per capita (current US)
https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

7. The World Bank (2022). Net Migration
https://data.worldbank.org/indicator/SM.POP.NETM?end=2023&start=2023&view=map

Annexes

Figure 1: Pairwise scatterplots n between all the variables included in the analysis

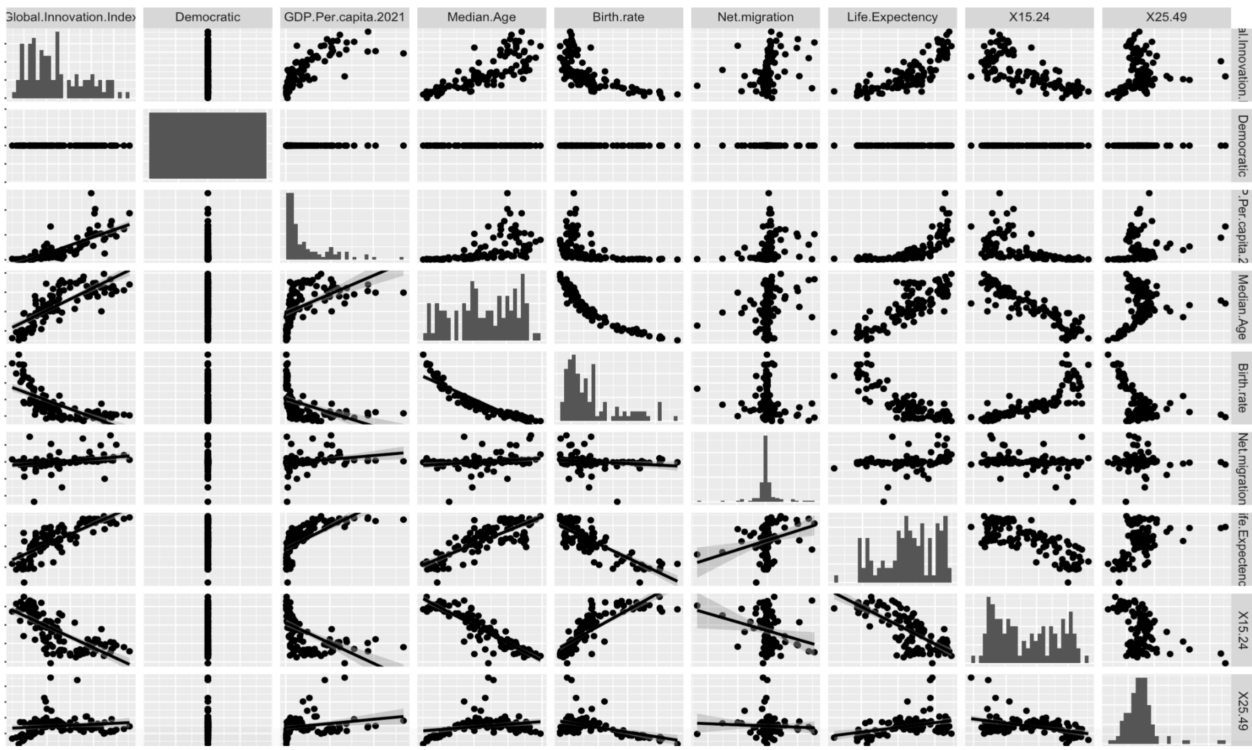


Table 1. Correlations among the selected variables

	Global.Innovation.Index	Democratic	GDP.Per.capita.2021	Median.Age	Birth.rate	Net.migration	Life.Expectancy	Group.15.24	Group.25.49
Global.Innovation.Index	1	0.205	0.791	0.783	-0.62	0.295	0.812	-0.729	0.148
Democratic	0.205	1	0.110	0.207	-0.096	0.284	0.037	-0.008	-0.376
GDP.Per.capita.2021	0.791	0.110	1	0.539	-0.459	0.258	0.713	-0.571	0.230
Median.Age	0.783	0.207	0.539	1	-0.924	0.252	0.815	-0.905	0.258
Birth.rate	-0.692	-0.096	-0.459	-0.924	1	-0.176	-0.818	0.823	-0.463
Net.migration	0.295	0.284	0.258	0.252	-0.176	1	0.270	-0.212	-0.054
Life.Expectancy	0.812	0.037	0.713	0.815	-0.818	0.270	1	-0.770	0.377
Group.15.24	-0.729	-0.008	-0.571	-0.905	0.823	-0.212	-0.770	1	-0.438
Group.25.49	0.148	-0.376	0.230	0.258	-0.463	-0.054	0.377	-0.438	1

Table 2. The initial model

Dependent variable: Global Innovation Index	Coefficient	Standard Error
Democratic	**0.640**	**(1.541)**
GDP Per capita 2021	**0.0002*** **	**(0.00003)**
Median Age	**0.297**	**(0.321)**
Birth rate	**-0.139**	**(0.230)**
Net migration	**0.00000**	**(0.00001)**
Life Expectancy	**0.348** **	**(0.162)**
Group 15/24	**-0.287**	**(0.439)**
Group 25/49	**-0.356** **	**(0.153)**
Constant	**13.188**	**(26.381)**

Notes: Observations: 112 R-squared: 0.827 Adjusted R-squared: 0.813 Residual Standard Error: 5.351 (df = 103) F Statistic: 61.462*** (df = 8; 103) Note: *p<0.1; **p<0.05; ***p<0.01

Table 3. The final model

Variable	***Coefficient***	***Standard Error***
GDP Per capita 2021	*0.001****	*(0.0001)*
Group 25/49	*0.220*	*(0.159)*
Median Age	*0.572****	*(0.074)*
*GDP Per capita 2021 * Group 25/49*	*-0.00001****	*(0.00000)*
Constant	*2.526*	*(4.701)*

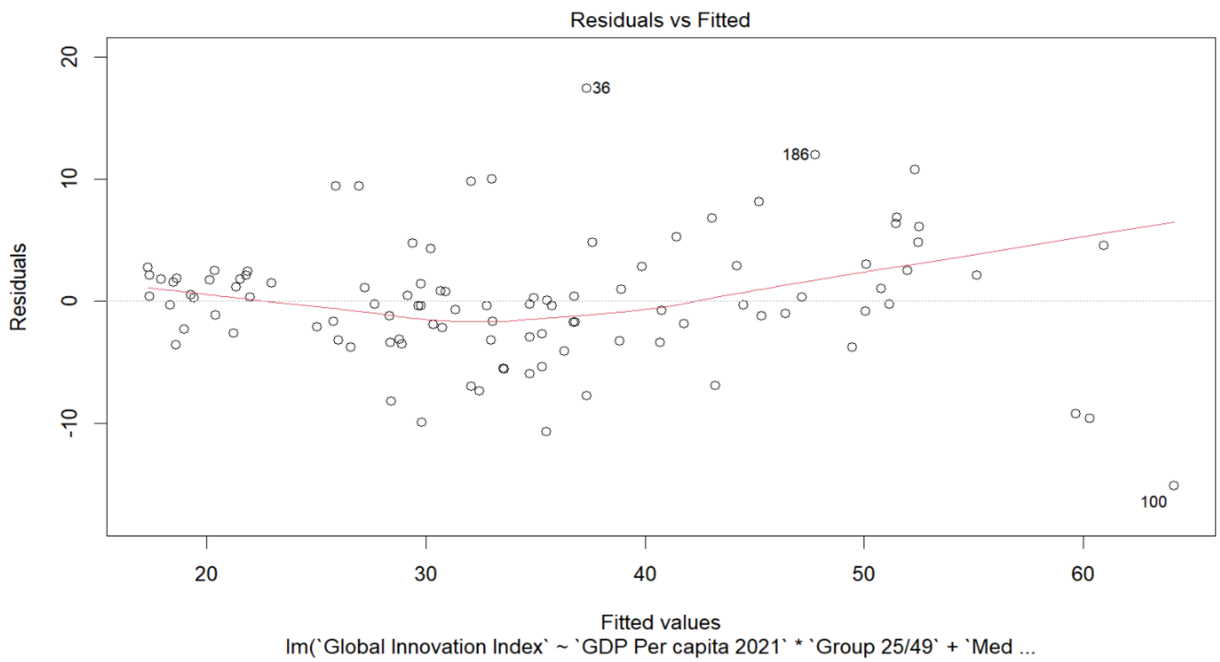
Notes: Observations: 112 R-squared: 0.832 Adjusted R-squared: 0.826 Residual Std. Error: 5.164 (df = 107) F Statistic: 132.887*** (df = 4; 107) Note: *p<0.1; **p<0.05; ***p<0.01

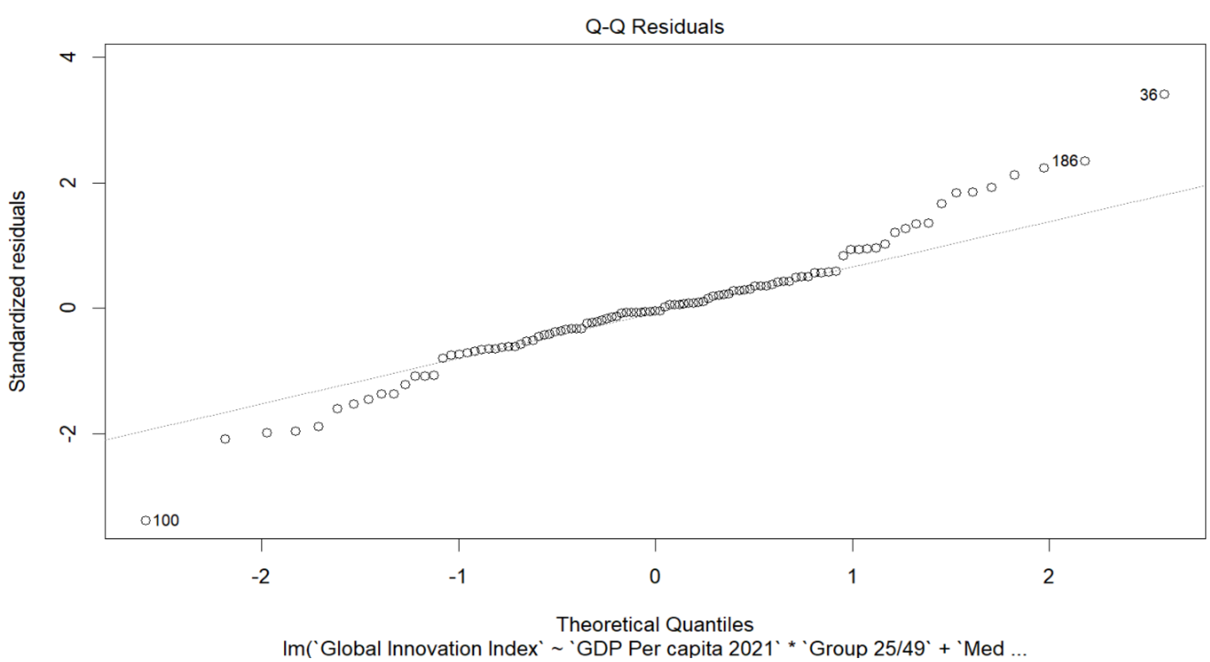
Table 4. The model after reducing unusual observations

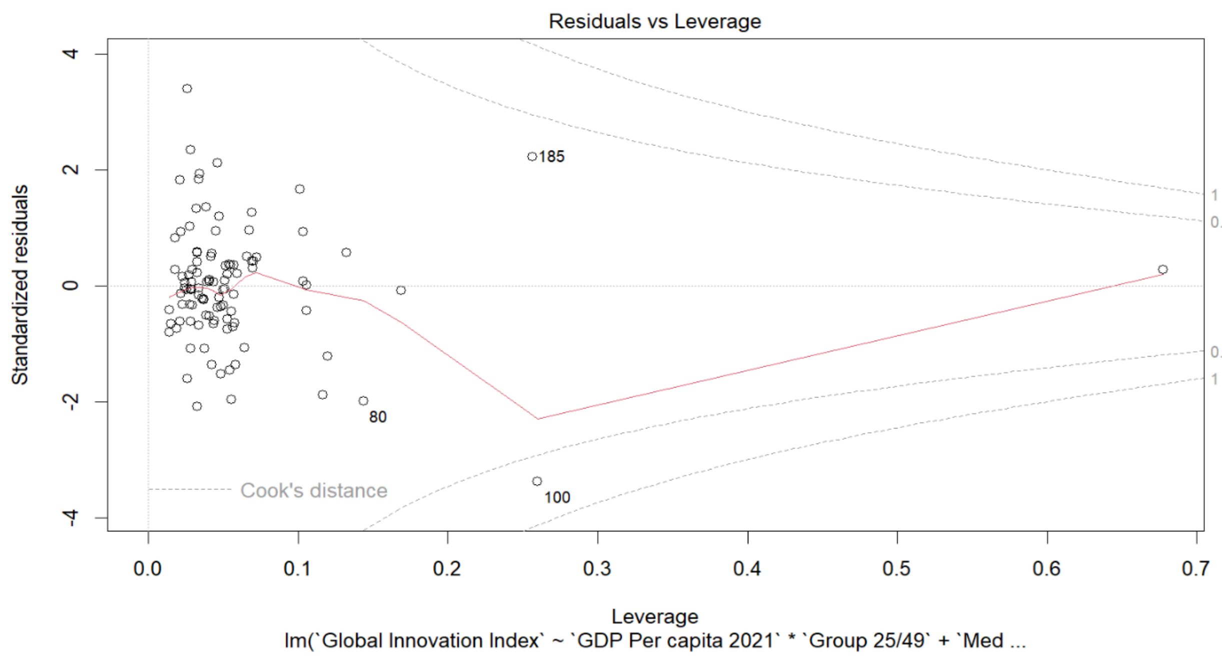
Variable	Coefficient	Standard Error
GDP Per capita 2021	**0.001*** **	**(0.0002)**
Group 25/49	**0.073**	**(0.182)**
Median Age	**0.458*** **	**(0.097)**
Life Expectency	**0.302* **	**(0.161)**
GDP Per capita 2021 * Group 25/49	**-0.00001** **	**(0.00000)**
Constant	**-10.030**	**(8.127)**

Notes: Observations: 103 R-squared: 0.835 Adjusted R-squared: 0.827 Residual Std. Error: 5.206 (df = 97) F Statistic: 98.324*** (df = 5; 97) Note: *p<0.1; **p<0.05; ***p<0.01

Figure 2. Plots for model with unusual observations






Residuals vs Leverage
Standardized residuals
4
2
0
-2
-4
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Leverage
lm(`Global Innovation Index` ~ `GDP Per capita 2021` * `Group 25/49` + `Med ...
185
80
100
Cook's distance
1
0.5
0.5
1

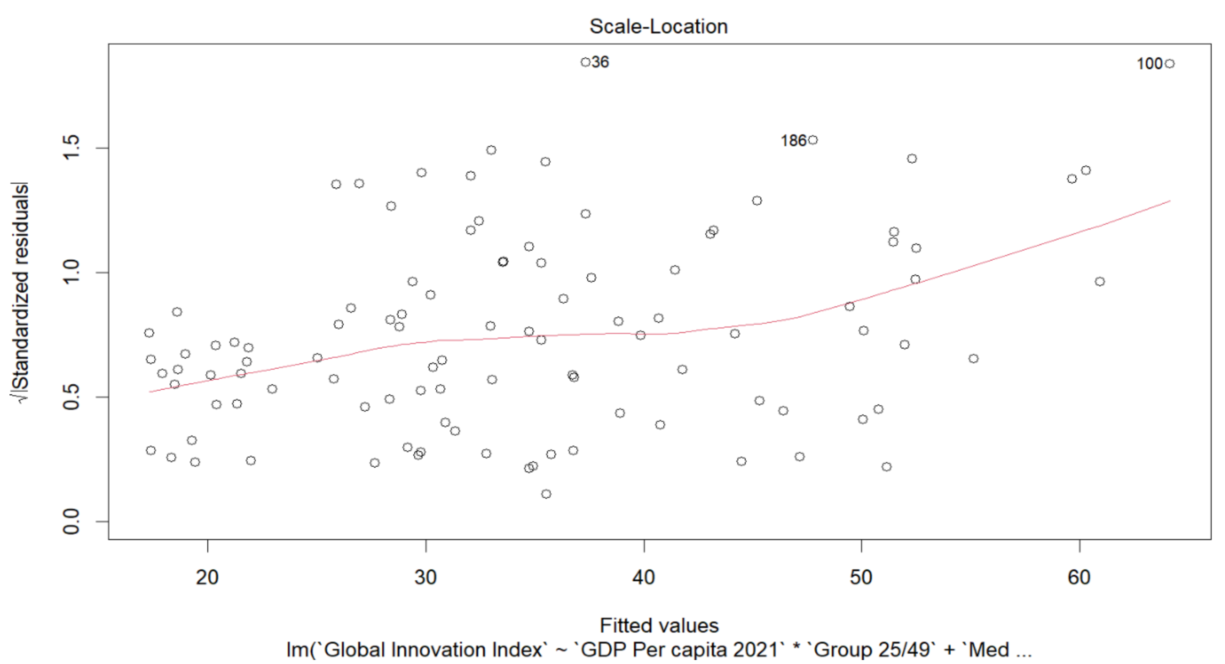

Scale-Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
20
30
40
50
60
Fitted values
lm(`Global Innovation Index` ~ `GDP Per capita 2021` * `Group 25/49` + `Med ...
36
186
100

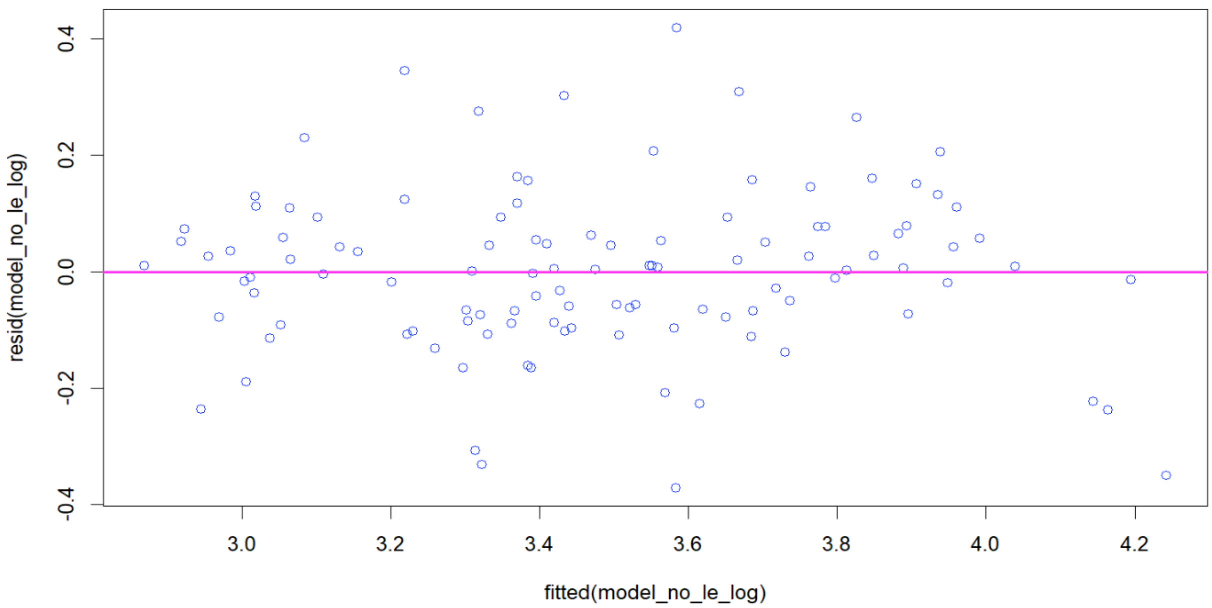
Table 5. VIFs

GDP Per capita 2021	Group 25/49	Median Age	GDP Per capita 2021 * Group 25/49
55.376	3.880	1.962	57.619

Table 6. Model with corrections

Variable	Coefficient	Standard Error
GDP Per capita 2021	0.00002***	(0.00000)
Group 25/49	0.012***	(0.004)
Median Age	0.019***	(0.002)
GDP Per capita 2021 * Group 25/49	-0.00000***	(0.00000)
Constant	2.281***	(0.130)

Figure 3. Fitted vs Residuals Plot and the Q_Q plot



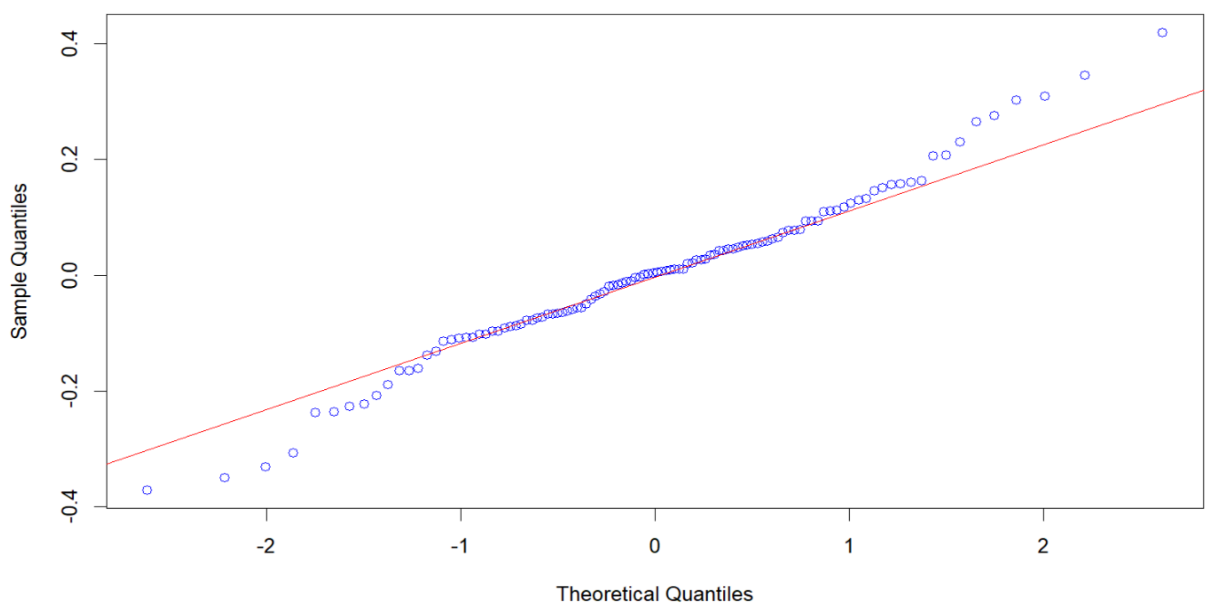


Figure 4. Histogram for normal distribution

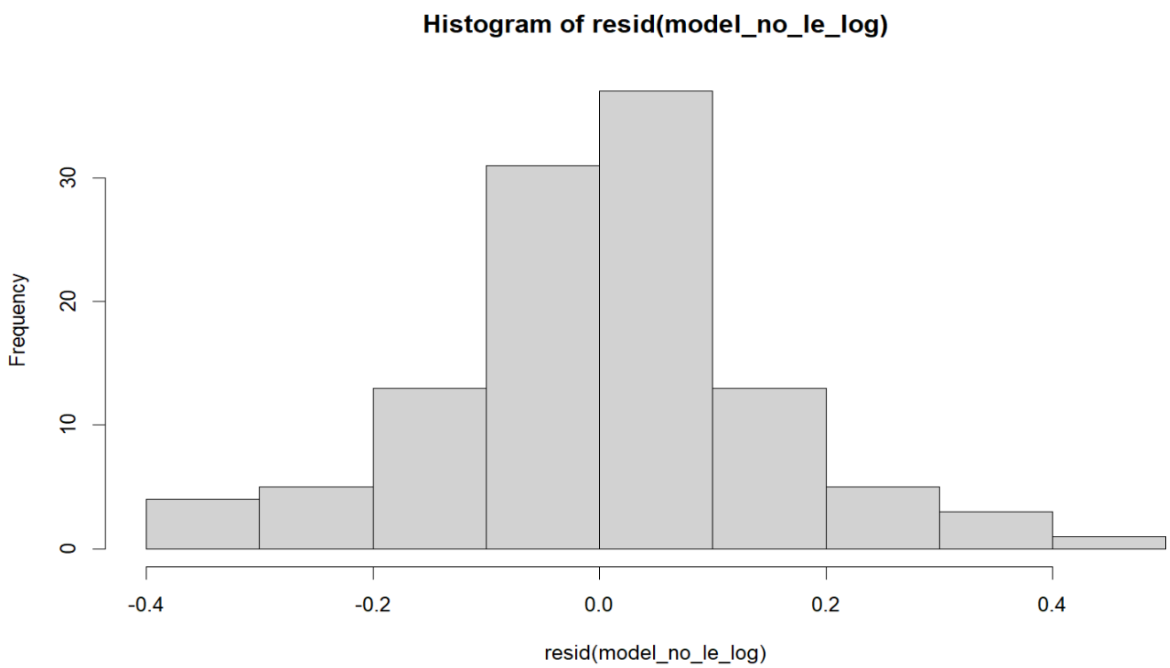


Table 7. Anova test

Statistic	N	Mean	St. Dev.	Min	Max
Df	9	12.333	34.000	1	103
Sum Sq	9	1,891.860	3,328.994	5.562	10,175.190
Mean Sq	9	1,567.375	3,355.296	5.562	10,175.190
F value	8	61.462	123.416	0.194	355.390
Pr(> F)	8	0.203	0.281	0.000	0.660