Incidence of Multicollinearity and Possible Violation of Weak Exogeneity Principle - Some Corrective measures and Application in a Recent Study

Nilay Tiwari, Roll no. 160454
Under the supervision of Dr. Sharmistha Mitra,
For SURGE Programme, IIT Kanpur, 2018
SURGE Roll no. S180086

Abstract

The Okun's Law proposed by Arthur Melvin Okun in 1962 establishes an emperical relationship between unemployment and GDP, whereby a percentage increase in unemployment causes approximately 2% fall in GDP growth. A recent paper[1] entitled 'Economic Growth Income Inequality: A revised cross-sectional econometric analysis of the global impact of income inequality on economic growth around the world' published in 2016 in Econometric Analysis U.G. Research by A. Hunter, W. Martinez, U. Patel of Georgia Institute of Technology under the supervision of Dr. S. Dhongde, explored this relationship.

This paper[1] examines the effects of inequality on GDP per capita growth in the year 2007 for 74 countries, around the world with the World Bank data and studies its relationship with other impactful factors on economic growth like Gross Savings Rate, Fertility Rate besides the Unemployment rate. The paper[1] uses GINI coefficient as a measure of the income inequality. The results show counterproductive relationship between unemployment and economic growth with respect to the Okun's Law, which warrants further research.

In this study, we look into the reasons that led to the positive relationship between unemployment and GDP per capita growth as opposed to the desired. The possible causes we probed6 include MultiCollinearity among the covariates and violation of the weak exogeneity assumption. The new approach, methodology and the corrective measures applied thereafter reiterates that Okun's Law still holds with respect to the current scenario data in some of the developed and developing economies, which were part of the original study.

The Results from the work will give us a better understanding on the variation of the dependence between GDP growth with Fertility Rate and Unemployment Rate with the change in the Economy from a developing to developed one.

Keywords: MutiCollinearity, Okun's Law, Panel Data Regression, Exogeneity.

1. Introduction

The Okun's Law proposed by Arthur Melvin Okun in 1962 establishes a relationship between unemployment and GDP growth rate, whereby a percentage increase in unemployment causes approximately 2% fall in GDP growth rate. Okun's Law was the first formulation of the relationship between unemployment and economic growth. Although a significant amount of literature is available on the above relationship however there is not a comparable amount of literature on the change in the relationship between the above two factors as the economy changes from developing to a developed one.

The situations for unemployed people in developed and developing countries are quite contrary. In most of the developed countries, there is a benefit given to the unemployed population. For example in the U.S. the Department of Labour[2] says 'The Department of Labor's Unemployment Insurance (UI) programs provide unemployment benefits to eligible workers who become unemployed through no fault of their own, and meet certain other eligibility requirements.The following resources provide information about who is eligible for these benefits and how to file a claim.'

Since, the developing countries have no such insurance or benefits plans for the unemployed population, the impact of the unemployed population in the developed countries is expected to be of a smaller magnitude compared to the developing countries. Therefore, there it is important to look into the relationship between economic growth and unemployment separately for developed and developing countries.

The relation between fertility rate and GDP growth has been a topic of great research. The majority of the literature points towards the initial increase of fertility rate with economic growth and the subsequent decline in fertility rate after the economic growth. The paper[3] by Yujie points out to the fact that 'no matter whether it is a rich or poor country, the effects are largely in the same direction, but they vary in their magnitude. This can be seen in the example: when rich countries receive greater influence than poor countries, with the influence arising capital forms of economic growth on the fertility rate. It shows that the negative impact that economic growth brings to the fertility rate is greater in rich countries than in poor countries'

The paper[4] by Gary S. Becker, Kevin M. Murphy, and Robert Tamura showed that 'When human capital is abundant, rates of return on human capital investments are high relative to rates of return on children, whereas when human capital is scarce, rates of return on human capital are low relative to those on children. As a result, societies with limited human capital choose large families and invest little in each member; those with abundant human capital do the opposite. '

These results point out to the fact that the dependence of the fertility rate with GDP growth might depend strongly on the fact if the country is a developing or a developed one.

A recent paper[1] entitled 'Economic Growth Income Inequality: A revised cross-sectional econometric analysis of the global impact of income inequality on economic growth around the world' published in 2016 in Econometric Analysis U.G. Research by A. Hunter, W. Martinez, U. Patel of Georgia Institute of Technology under the supervision of Dr. S. Dhongde

This paper[1] examines the effects of inequality on GDP per-capita growth in the year 2007 for 74 countries, both developed and developing ones around the world with the World Bank data and compares that relationship to other impactful factors on economic growth like Gross Savings Rate, Fertility Rate, and the Unemployment rate. The paper[1] uses GINI coefficient as a measure of the income inequality. The results shows positive relationship between unemployment and economic growth, quite contrary to Okun's Law. The results also show positive relation of economic growth between GDP per capita growth in the year 2007 with GINI, Gross savings rate and fertility rate.

The study by Hunter et. al[1] concludes with the remark that this counter productive relationship could be the result of the technical factors involved in the regression modelling approach. These are 1) The presence of multicollinearity among the exogenous variables in the model or, 2) Violation of the weak exogeneity assumption of these variables with the error term in the model.

It is well known that multicollinearity not only causes abnormally high variance of the estimated coefficients, but also abnormally high magnitude of these coefficients. We believe that the problem of multicollinearty produces wrong signs for the coefficients of the independent variables of the model. The multicollinearity between the covariates can be identified with the help of the Variance Inflation Factor and if required, some of the covariates need to be dropped with the method of Subset Regression Models and Variance Decomposition.

To take the case of the possible shortcomings of the study we plan to apply the Variance Inflation Factor. We also propose that the Fertility Rate is a factor which depends upon the fact weather the country is a Developed or a Developing Economy. We plan on dropping the Correlated variables by the method of Variance Inflation Factor and then constructing the subsequent Subset Regression Models for all countries. A Panel Data Regression for the Models of 11 countries in the 15 year range 2002-2016 is also constructed.

This study differs from the approach of the previous paper[1] in the sense that here time series data(annual) has been used in the model. 11 countries for the period of 2002 to 2016 and the various countries make 11 models which include 4 developed and 7 developing countries. The measure used in this study for economic growth is also opposed to the one used in the previous paper[1]. We have used the usual GDP growth rate which is considered to be a standard measure of economic growth of a country. The other contribution of this work is to indicate the opposing relationship of the fertility rate and economic growth with the change in the economy from a developed to a developing one.

The paper proceeds in the following order: Section 2 and 3 contains the relevant theory in this area and the available literature in the area. Section 4 describes the Data used by us in our simulations. Section 5 and 6 consist of the results produced by us and the conclusions are elaborated in section 7. In the end Appendix provides the detailed output we have used to summarize our results.

2. Multicollinearity in Multiple Linear Regression Model

2.1. Multiple Linear Regression Model

A regression model that involves more than one regressor variable is called a multiple regression model. For Ref. *Chapter 3, Multiple Linear Regression, Introduction to Linear Regression Analysis*[5]. In general, the response y may be related to k regressor or predictor variables.
The model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \epsilon \tag{1}$$

is called a mnltiple linear regression model with k regressors. The parameters β_j , j = 0, 1,...., k , are called the regression coefficients . This model describes a hyperplane in the k - dimensional space of the regressor variables x_j . The parameter β_j represents the expected change in the response y per unit change in x_j when all of the remaining regressor variables $x_{j'}$ ($j' \neq$ j) are held constant . It is more convenient to deal with multiple regression models if they are expressed in matrix notation. This allows a very compact display of the model, data, and results. In matrix notation, the model given by Eq. (2) is

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon} \tag{2}$$

where

$$\mathbf{y} = \begin{pmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \\ \mathrm{y}_3 \\ . \\ . \\ . \\ \mathrm{y}_k \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & \mathrm{x}_{11} & \mathrm{x}_{12} & \ldots & \mathrm{x}_{1k} \\ 1 & \mathrm{x}_{21} & \mathrm{x}_{22} & \ldots & \mathrm{x}_{2k} \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 1 & \mathrm{x}_{n1} & \mathrm{x}_{n2} & \ldots & \mathrm{x}_{nk} \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_k \end{pmatrix} \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ . \\ . \\ . \\ \epsilon_k \end{pmatrix}$$

In general, $\mathbf{y}$ is an n x 1 vector of the responses, $\mathbf{X}$ is an n x $\overline{k+1}$ matrix of the levels of the regressor variables, $\boldsymbol{\beta}$ is a $\overline{k+1}$ x 1 vector of the regression coefficients, and $\boldsymbol{\epsilon}$ is an n x 1 vector of random errors.

2.2. The Ordinary Least Squares(OLS) Estimator

The method of least squares is used to estimate the regression coefficients in eq. (2) . Suppose that n observations are available, and let y_i denote the i^{th} observed response and x_{ij} denote the i^{th} observation or level of regressor x_j . The data will appear as in Table 3.1 . We assume that the error term ϵ in the model has $E(\epsilon) = 0$, $Var(\epsilon) = \sigma^2$, and that the errors are uncorrelated.
We wish to find the vector of least squares estimators, $\hat{\boldsymbol{\beta}}$, that minimizes the mean square error term

$$S(\boldsymbol{\beta})=\sum_{i=1}^{n} \epsilon_i^2 = \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

which is given by,

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{3}$$

provided that the inverse matrix $(\mathbf{X}'\mathbf{X})^{-1}$ exists. The $(\mathbf{X}'\mathbf{X})^{-1}$ matrix will always exist if the regressors are linearly independent , that is, if no column of the X matrix is a linear combination of the other columns.

The estimators in (3) are the OLS estimators, which are the Best Linear Unbiased Estimators(BLUEs) of the coefficients. For Ref. *Chapter 3, Multiple Linear Regression, Introduction to Linear Regression Analysis*[5].

2.3. Multicollinearity

If there is no linear relationship between the regressors, they are said to be orthogonal. However, in some situations the regressors are nearly linearly related, and in such cases the inferences based on the regression model can be misleading or erroneous. When there are near - linear dependencies among the regressors, the problem of multicollinearity is said to exist. For Ref. *Chapter 9, Multicollinearity, Linear Regression Analysis*

2.4. Sources of Multicollinearity

When both regressors and responses have been standardized, the $\mathbf{X}'\mathbf{X}$ is a $\overline{k+1}$ x $\overline{k+1}$ matrix of correlations between the regressors and $\mathbf{X}'\mathbf{y}$ is a $\overline{k+1}$ x 1 vector of correlations between the regressors and the response. Let the j^{th} column of the $\mathbf{X}$ matrix be denoted $\mathbf{X_j}$, so that $\mathbf{X} = [\ \mathbf{X_1}\ ,\ \mathbf{X_2}\ ,\ .\ .\ .\ .\ ,\ \mathbf{X_p}\ ]$. Thus, $\mathbf{X_j}$ contains the n levels of the j^{th} regressor variable. We may formally define multicollinearity in terms of the linear dependence of the columns of X . The vectors $\mathbf{X_1}$, $\mathbf{X_2}$, ... , $\mathbf{X_p}$ are linearly dependent if there is a set of constants t_1 , t_2 , ... , t_p , not all zero, such that

$$\sum_{j=1}^{p} t_j \mathbf{X_j} = 0 \tag{4}$$

If Eq. (4) holds exactly for a subset of the columns of X , then the rank of the $\mathbf{X'X}$ matrix is less than k+1 and $(\mathbf{X'X})^{-1}$ does not exist. However, suppose that Eq. (9.1) is approximately true for some subset of the columns of X . Then there will be a near - linear dependency in $\mathbf{X'X}$ and the problem of multicollinearity is said to exist. Note that multicollinearity is a form of ill - conditioning in the $\mathbf{X'X}$ matrix. Furthermore, the problem is one of degree, that is, every data set will suffer from multicollinearity to some extent unless the columns of $\mathbf{X}$ are orthogonal.

There are four primary sources of multicollinearity :
1. Data collection method (mostly in case of survey data)
2. Constraints imposed on the model coefficients
3. An over specified model

2.5. Effects of Multicollinearity and Variance Inflation Factor

For the case of Multiple Regressors, multicollinearity produces similar effects. It can be shown that when both regressors and responses have been standardized, the diagonal elements of the $\mathbf{C} = (\mathbf{X'X})^{-1}$ matrix are

$$C_{jj} = \frac{1}{(1 - R_j^2)}, j = 1, 2, ...p \tag{5}$$

where R_j^2 is the coefficient of multiple determination from the regression of x_j on the remaining k regressor variables. If there is strong multicollinearity between x_j and any subset of the other k, regressors, then the value of R_j^2 will be close to unity. Since the variance of β_j is $\mathrm{Var}(\beta_j) = C_{jj}\sigma^2 = (1 - R_j^2)^{-1}\sigma^2$, strong multicollinearity implies that the variance of the least - squares estimate of the regression coefficient β_j is very large

Multicollinearity also tends to produce least - squares estimates $\hat{\beta}$ that are too large in absolute value. For Ref. *Multicollinearity, Econometric Analysis*[6]. To see this, consider the squared distance from $\hat{\boldsymbol{\beta}}$ to the true parameter vector $\boldsymbol{\beta}$, for example,

$$L_1^2 = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$$

Where expected squred distance, $\mathrm{E}(L_1^2)$ is

$$\begin{aligned}\mathrm{E}(L_1^2) = \mathrm{E}((\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})) = \textstyle\sum_{j=1}^{p} E((\hat{\beta} - \beta))^2 \\ = \textstyle\sum_{j=1}^{p} Var(\hat{\beta}_j) = \sigma^2 Tr(\mathbf{X}'\mathbf{X})\end{aligned}$$

Which gives,

$$E(L_1^2) = \sigma^2 Tr(\mathbf{X}'\mathbf{X}^{-1}) \tag{6}$$

where the trace of a matrix (abbreviated Tr) is just the sum of the main diagonal elements. When there is multicollinearity present, some or more of the eigenvalues of $\mathbf{X}'\mathbf{X}$ will be close to zero. Since the trace of a matrix is also equal to the sum of its eigenvalues, Eq. (7) becomes

$$E(L_1^2) = \sigma^2 \sum_{j=1}^{p} \frac{1}{\lambda_j} \tag{7}$$

Indicating an abnormally large expected value of the sum of sq. of the distance between the estimated coefficients and their true values.

If x_j is nearly orthogonal to the remaining regressors, R_j^2 is small and C_{jj} is close to unity, while if x_j is nearly linearly dependent on some subset of the remaining regressors, R_j^2 is near unity and C_{jj} is large. Since the variance of the j th regression coeffi cients is $C_{jj}\sigma^2$, we can view C_{jj} as the factor by which the variance of β_j is increased due to near - linear dependences among the regressors.
We call,

$$VIF_j = C_{jj} = (1 - R_j^2)^{-1} \tag{8}$$

the variance inflation factor. As a rule of thumb, if any of the VIFs exceeds 5 or 10, it is an indication that the associated regression coefficients are inefficiently estimated because of multicollinearity.

3. Panel Data Regression

Cross-sectional data is the data collected on several individuals/units at one point in time. Time series data data collected on one individual/unit over several time periods. A regression model combining these two data is done by Panel Data Regression. Panel data are repeated cross-sections over time. For detailed analysis of the below results, Ref. to *Part IV, Panel Data, Econometric Analysis*[6]. Broadly, there are two types of model for panel data, Fixed and Random Effects Models.

The various models for Panel Data are as follows :

3.1. Pooled Estimator

This is the simplest estimator for the Panel Data It essentially ignores the panel structure of the data and uses OLS assuming the classical assumptions of OLS estimator. It is the easiest model available but is essentially not suitable for our use as it ignores the panel structure of the data.

3.2. Fixed Effects Model

When using FE we assume that something within the individual may impact or bias the predictor or outcome variables and we need to control for this. This is the rationale behind the assumption of the correlation between entitys error term and predictor variables. FE remove the effect of those time-invariant characteristics so we can assess the net effect of the predictors on the outcome variable. Another important assumption of the FE model is that those time-invariant characteristics are unique to the individual and should not be correlated with other individual characteristics. Each entity is different therefore the entitys error term and the constant (which captures individual characteristics) should not be correlated with the others. If the error terms are correlated, then FE is no suitable since inferences may not be correct and you need to model that relationship (probably using random-effects), this is the main rationale for the Hausman test (presented later on in this document).
The equation for the fixed effects model becomes:

$$Y_{it} = \beta_i X_{it} + \alpha_i + \epsilon_{it} \tag{9}$$

here
- α_i is the unknown intercept for each entity (n entity-specific intercepts).
- Y_{it} is the dependent variable where i = entity and t = time.
- X_{it} represents one independent variable (IV),

- β_i is the coefficient for that IV,
- ϵ_{it} is the error term.

The fixed-effects model controls for all time-invariant differences between the individuals, so the estimated coefficients of the fixed-effects models cannot be biased because of omitted time-invariant characteristics

3.3. Random Effects Model

The rationale behind random effects model is that, unlike the fixed effects model, the variation across entities is assumed to be random and uncorrelated with the predictor or independent variables included in the model. An advantage of random effects is that you can include time invariant variables. In the fixed effects model these variables are absorbed by the intercept. Random effects assume that the entitys error term is not correlated with the predictors which allows for time-invariant variables to play a role as explanatory variables. In random-effects you need to specify those individual characteristics that may or may not influence the predictor variables. The problem with this is that some variables may not be available therefore leading to omitted variable bias in the model.

The random effects model is:

$$Y_{it} =_{it} +\epsilon_{it} \tag{10}$$

$$\epsilon_{it} = \alpha_i + \eta_{it} \tag{11}$$

3.4. Fixed Effects vs Random Effects Model

To decide between fixed or random effects you can run a Hausman test where the null hypothesis is that the preferred model is random effects vs. the alternative the fixed effects (see Green, 2008, chapter 9). It basically tests whether the unique errors ϵ_i are correlated with the regressors, the null hypothesis is they are not.

4. Data

Our paper analyzes the cause of the wrong relationship between unemployment rate and GDP growth as shown by the previous paper[1] which used the cross sectional data for 74 countries and did a regression for GDP growth as the dependent variable and fertility rate, unemployment rate, gross savings rate and GINI coefficient as the independent variables. The data used in their exercise for the year 2007.

But the cross- sectional regression on the 74 countries is not suitable for showing the correct relationship. Their approach, at best will be able to study the presence of multi-collinearity in the explanatory variables for the different countries. This is logically a very weak finding without any significance especially, when the various countries are not from the same stage of economic development and most importantly the data is only for a single point of time.

A better way is to use the time series data instead and form regression models for different countries and compare the results for all countries and then analyze the data and find possible patterns in the relationships of developed and developing countries. The previous paper[1] used the percent change of GDP as our dependent variable (y) and decided on a range of five years, 2007-2012, in order to capture the short and medium economic growth rate.

i.e. the measure for GDP growth in year 2007 used was

$$\text{GDP growth}(\%) = \frac{(GDPpercapita(year-2012)-GDPpercapita(year-2007))}{GDPpercapita(year-2007)} * 100\%$$

The measure used in our work for GDP growth is the World Bank data for annual GDP growth rate(%). The Rest of the data consists of fertility rate, unemployment rate, gross savings rate and GINI coefficient for 11 Countries, consisting of 4 developed and 7 developing countries for a period of 15 years from 2002 to 2016, including both. There was no missing data and the 15 year range from 2002-2016 remains descriptive of the relationships between the regressors

The countries have been classified as developing and developed countries.These two categories were created using the World Bank classification system: Within the data-set, countries which have an income of $12,736 or more were classified as the developed, while a zero is placed for countries which have an income lower than $12,736 were classified as the developing countries.

4.1. Previous Results

The previous paper[1] results are summed here. Overall, the OLS regression models showed that inequality and economic growth share a positive relationship. By obtaining data from the World Bank, sample size of 74 countries were constructed consisting of both developed and developing countries and the utilization of independent variables such as the GINI coefficient, gross savings rate, unemployment, and fertility to measure economic growth (change in GDP per capita for the period 2007-2012). In addition to the GINI coefficient, the analysis showed that gross savings rate, unemployment, and fertility all possessed positive impacts on economic growth for the time period of 2007 to 2012. Although the dependence of unemployment was insignificant, all other variables had significant impact on the economic growth.

5. Results

The results of the Linear Regression Model :

$$gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \beta_4 fer + \epsilon \tag{12}$$

for the 11 countries have been shown in detail in the Appendix.
The value of the coefficients of fertility rate and unemployment rate in the above regression model for the 11 countries along with the Class(high income or low income) is as follow :

Country No.	Coeff. for unemp	Coeff. for fer	Economy
1	-1.1872	-51.527	High Income
2	7.9485	-17.0472	High Income
3	-0.182	1.8242	Low Income
4	-1.34	2.974	Low Income
5	-2.25	-14.699	High Income
6	0.85	-2.084	High Income
7	-1.252	-3.683	Low Income
8	2.057	89.04	Low Income
9	-1.462	-0.2495	Low Income
10	3.263	8.8507	Low Income
11	-2.209	-1.038	Low Income

Thus, from the 4 developed countries, number of countries with negative coefficient of unemployment rate is 2, number of countries with negative coefficient of fertility rate is 4, whereas for 7 developing countries, number of countries with negative coefficient of unemployment rate is 5,number of countries with negative coefficient of fertility rate is 3.

5.1. Multicollinearity Analysis

It is necessary to perform a Multicollinearity to be sure that the OLS carried is the unbiased estimator of the Model Coefficients. For the Multicollinearity, the Varinace Inflation Factor was calculated for all regressors for all 11 Models. The detailed results are as follows,

Country No.	VIF for unemp	VIF for gsavs	VIF for GINI	VIF for fer
1	3.715	3.767	2.617	2.657
2	2.902	1.266	1.986	2.500
3	3.595	1.804	1.252	2.994
4	1.097	1.319	6.361	6.063
5	5.696	3.2556	2.972	1.204
6	1.152	1.653	2.726	2.602
7	2.156	1.020	2.223	3.401
8	4.183	4.415	5.210	2.249
9	9.270	4.570	13.586	6.441
10	39.399	4.224	36.120	17.341
11	6.549	3.136	5.945	6.583

As we know that the VIF < 5 means that the data is free of Multicollinearity and the analysis shows unbiased results but we can clearly see that the VIF for unemployment rate and fertility rate exceed 5 for 4 countries whereas for GINI coefficient, it exceeds 5 for 5 countries. Gross savings on the other hand has VIF < 5 for all 11 countries. This warrants a possible relationship present between unemployment rate, GINI coefficient and fertility rate. Since, these 3 can be correlated in only one manner i.e. when all 3 are correlated together which calls for the Subset regression to identify the Most correlated regressor and to drop that regressor.

5.2. Subset Regression

In order to identify the most suitable regressor to be dropped from the Model, we carried out the Subset Regression of all the the correlated regressors. In the subsequent linear regressions, the correlated regressors were made the dependent variables and the other variables as the independent ones.

Three models were constructed as follows:

Model 1 with GINI coefficient as the dependent variable :

$$GINI = \beta_0 + \beta_1 unemp + \beta_2 fer + \epsilon \tag{13}$$

Model 2 with fetility rate as the dependent variable :

$$fer = \beta_0 + \beta_1 GINI + \beta_2 unemp + \epsilon \tag{14}$$

Model 3 with unemployment rate from the initial model:

$$unemp = \beta_0 + \beta_1 fer + \beta_2 GINI + \epsilon \tag{15}$$

The value of adjusted R^2 statistic for the three models alongwith that of the initial model is as follows :

Country No.	GINI	fer	unemp
1	.5536	.3092	.4721
2	0.3122	.5242	.5874
3	.04172	.4939	.4233
4	.8078	.8076	-0.06971
5	.5975	.000396	.5423
6	.5406	.536	-0.1485
7	.471	.6552	.4503
8	.4066	-0.1094	.387
9	.9122	.7867	.83
10	.9419	.9327	.9074
11	.7452	.7813	.758

Since the value of the adjusted R^2 is the highest for the Model for the subset regression with GINI coefficient as the dependent variable for 7 out of 11 countries, it is clear that the GINI coefficient is the one best explained by the other two correlated variables, i.e., unemployment rate and the fertility rate. So, it is reasonable to drop GINI coefficient from our initial Model to solve the problem of Multicollinearity.

In order to further check our results for the most suitable regressor to be dropped from the Model, we carried regressions with all of the the correlated regressors dropped from our initial model one by one and then finding the value of the adjusted R^2 and comparing the results with our initial model. Three models were constructed as follows:

Model 1 with GINI coefficient dropped from the initial model:

$$gdpchange = \beta_0 + \beta_1 unemp + \beta_2 gsavs + \beta_3 fer + \epsilon \tag{16}$$

Model 2 with fertility rate dropped from the initial model:

$$gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \epsilon \tag{17}$$

Model 3 with unemployment rate dropped from the initial model:

$$gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 fer + \epsilon \tag{18}$$

The value of the adjusted R^2 statistic for the three models along with that of our initial model is as follows :

Country No.	initial model	GINI dropped	fer dropped	unemp dropped
1	0.418	0.5284	0.5064	0.3698
2	.7338	0.7579	0.5445	0.7086
3	.1329	0.001169	0.2057	0.2019
4	-0.09693	-0.02394	-0.0001248	-0.129
5	.5327	0.1485	0.209	0.3528
6	.5602	0.3912	0.5079	0.4588
7	-.1859	-0.08446	-0.1109	-0.1882
8	-.01653	-0.01483	-0.033	-0.1077
9	.1454	0.1061	0.2231	-0.1806
10	.4506	0.4636	0.4837	0.4376
11	-.1203	-0.06712	-0.01912	-0.1545

Since the value of the adj. R^2 is the higher for the Model with the GINI coefficient dropped out than our initial model for of the the ones times, it is clear that the new model with GINI coefficient dropped will not only solve the problem of Multicollinearity but will also have better information in the form of the adjusted R^2 than the initial model.

In order to test if the Problem of Multicollinearity has been solved or not, by removing the GINI coefficient, we have evaluated the VIF matrix for the model with GINI coefficient dropped.

Country No.	VIF for unemp	VIF for gsavs	VIF for fer
1	2.419	3.762	2.050
2	2.403	1.081	2.485
3	3.485	1.754	2.536
4	1.094	1.254	1.306
5	3.022	3.039	1.088
6	1.151	1.540	1.371
7	1.548	1.012	1.535
8	1.051	1.666	1.688
9	4.229	4.470	3.241
10	17.320	2.349	13.169
11	6.535	2.415	4.306

Since, the VIF for all the regressors exceed only 3 times, one time for fertility rate and two times for unemployment rate, it is clear that the problem of Multicollinearity has been solved and the model has a better adjusted R^2 than the initial model.

5.3. *New Proposed Model*

The Model proposed by us is the one with the GINI coefficient dropped from initial model:

$$gdpchange = \beta_0 + \beta_1 fer + \beta_2 gsavs + \beta_3 unemp + \epsilon \tag{19}$$

So, the coefficients of the unemployment in the above model for 11 countries is as follows :

Country No.	Coeff. for unemp	Economy
1	-1.191	High Income
2	7.782	High Income
3	-0.328	Low Income
4	-1.369	High Income
5	-0.116	High Income
6	0.8326	Low Income
7	-1.092	Low Income
8	0.8074	Low Income
9	-0.8812	Low Income
10	1.391	Low Income
11	-2.15	Low Income

Thus, from the 4 developed countries, number of countries with negative coefficient of unemployment rate is 2, number of countries, whereas for 7 developing countries, number of countries with negative coefficient of unemployment rate is 5

The unemployment rate shows expected results, for the developing countries, a clear majority of the countries show a negative relationship between the unemployment rate and GDP growth rate as expected b the OKUN's Law. However no clear relationship is clear for the unemployment rate and GDP growth in the developed ones.

These results can be attributed to the fact that the developed countries have Unemployment Insurance(UI)[2] allocation for the unemployed population whereas the developing countries have no such grant available. Therefore, the unemployment has no such clear 'negative' relationship with the GDP growth in the developed ones. But for the case of he developing countries, the relationship is clear, in accordance with the OKUN's Law.

5.4. Analysis for Fertility Rate

The coefficients of the fertility rate in our proposed model for 11 countries are as follows :

Country No.	Coeff. for fer	Economy
1	-51.88	High Income
2	-17.015	High Income
3	-2.37	Low Income
4	-5.0027	High Income
5	-9.568	High Income
6	0.072	Low Income
7	-2.4709	Low Income
8	48.347	Low Income
9	32.099	Low Income
10	2.4267	Low Income
11	4.1428	Low Income

Thus, from the 4 developed countries, number of countries with negative coefficient of fertility rate is 3, whereas for 7 developing countries, number of countries with negative coefficient of fertility rate is also 3.

The results shown above clearly indicate the developed countries have a NEGATIVE relationship between fertility rate and GDP growth rate. It was expected as the previous literature[3] points out towards the fact that the Fertility Rate increases in the initial stage of the economic growth and then becomes stagnant as the economy reaches a stable stage and gradually begins to decrease with the country falling in the category of the developed ones.

So, it can be expected that the fertility rate will have negative impacts on the economy once the country has become a developed one. However for the developing countries, there is no clear relationship between the fertility rate and the GDP growth rate. The developing countries show both positive and negative relationship with the economic growth.

6. Panel Data Regression

Since the data available with us contains both time series and cross-sectional data, we can apply Panel Data Regression for the Panel Data available. The Panel Data Regression is carried out for the time series data for 15 years from 2002 to 2016 for 11 countries. For all the estimates below, n=11, T=15, N=165

6.1. With Pooled OLS Estimator

The statistics for the estimator are as follows :
P value for unemp : 0.785998
Coefficient for unemp = 0.021135

6.2. With Between Estimator

The statistics for the estimator are as follows :
P value for unemp = 0.003987
Coefficient for unemp = 0.183070

6.3. With Within Estimator of Fixed Effects Model

The statistics for the estimator are as follows :
Pvalue for unemp = 0.0004345
Coefficient for unemp = -0.612706

6.4. With Random Effects Estimator

The statistics for the estimator are as follows :
Pvalue for unemp = 0.785998
Coefficient for unemp = 0.021135

6.5. LM Test for Random Effects vs Pooled OLS

Model : $gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \epsilon$
normal = -1.1353, p-value = 0.8719
alternative hypothesis: significant effects alternative hypothesis : significant effects
So, Random Effects turns out to be better than Pooled Estimator

6.6. LM Test for Fixed Effects vs Pooled OLS

F test for individual effects

Model : $Gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \epsilon$

F = 2.2511, df1 = 10, df2 = 151, p-value = 0.01769

alternative hypothesis: significant effects

So, Fixed Effects turns out to be better than Pooled Estimator

6.7. LM Test for Fixed Effects vs Random Effects

Hausman test

Model : $Gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \epsilon$

chisq = 21.143, df = 3, p-value = 9.834e-05

alternative hypothesis: one model is inconsistent alternative hypothesis : one model is inconsistent

Since, one of the model is inconsistent, therefore Fixed Effects turns out to be better than Random Effects Estimator.

So, fixed effects Model is the best model for panel data regression.

The coefficient for the unemployment in the Fixed Effects Model is -0.612706, negative as expected from the Okun's law.

7. Conclusion

The work concludes that the reason for the opposed relationship between unemployment rate and GDP growth in the previous paper[1] Economic Growth Income Inequality: A revised cross-sectional econometric analysis of the global impact of income inequality on economic growth around the world published in Econometric Analysis U.G. Research by A. Hunter, W. Martinez, U. Patel of Georgia Institute of Technology by S. Dhongde was due to the strong Multicollinearity present in the data. The unemployment rate, fertility rate and GINI coefficient were strongly correlated.

On producing subset regression for the correlated regressors, we concluded that the GINI coefficient was best explained by the other two correlated regressors and had value of adj. R^2 higher than the initial model for 7 out of 11 countries. Therefore, to solve the problem of multicollinearity, we dropped GINI coefficient.

After removing the GINI coefficient, we found that the relationship between unemployment rate and GDP growth differs with the economy from a developed to a developing one. So, upon analyzing we found that the unemployment has a clear negative relationship with economic growth in developing countries and not a clear relationship in developed ones. This can be attributed to the fact that the developed countries have an Unemployment Insurance(UI) and other benefits for the unemployed population whereas the developing countries have no such grant which leads to the above relationship.

For the relationship between fertility rate and GDP growth, we found that the developed countries have a NEGATIVE relationship between fertility rate and GDP growth rate. It was expected as the previous literature[3] points out towards the fact that the Fertility Rate increases in the initial stage of the economic growth and then becomes stagnant as the economy reaches a stable stage and gradually begins to decrease with the country falling in the category of the developed ones. However for the developing countries, there is no clear relationship between the fertility rate and the GDP growth rate.

After performing the panel data regression for cross sectional data for 11 countries and the time series available for the years 2002-2016 for all the estimators, i.e. Pooled OLS, Fixed effects estimator and the Random Effects Estimator, we found that the fixed effects and random effects model were both better than pooled OLS estimator by LM Hausman Test an fixed effects to be better than random effects. The fixed effects model had Coefficient for unemp -0.612706 and p-value of 0.003987.

The aim of this study was to find the possible reasons for the counter productive relationship between unemployment rate and GDP growth rate with respect to the OKUN's law in a previous paper[1] which were probed to be the multicollinerity present in the regressors or the violation of the weak exogeneity principle. We found that there was strong multicollinearity present in the 3 regressors which were unemployment rate, fertility rate and GINI coefficient and upon performing subset regression for the 3 regressors, we found GINI coefficient to be most correlated variable and dropping th variable not only solved the problem of multicollinearity in the model but also caused the value of the adjusted R^2 to increase from the initial model. The other possible reason for the wrong relationship might be the violation of the weak exogeneity principle which we haven't checked in this study, however this possibility could be verified in a future work and could be easily solved using Instrumental Variable Estimation.

Due to the missing data, we had to go with just 11 countries. But the varying relationship between fertility rate, unemployment rate and economic growth for developed and developing countries is an interesting topic for further research. Better data and accountability, especially in regards to the GINI coefficient, would also help for further research and analysis.

References

[1] A. Hunter, W. Martinez and U. Patel(2016). *Economic Growth Income Inequality: A revised cross-sectional econometric analysis of the global impact of income inequality on economic growth around the world*(Econometric Analysis U.G. Research). Georgia Institute of Technology, Atlanta, USA. Retrieved from https://smartech.gatech.edu/handle/1853/54741.

[2] Unemployment Insurance *U.S. Department Of Labour*. LUND University, Lund, Sweden. Retrieved from https://www.dol.gov/general/topic/unemployment-insurance

[3] Yujie Li(2015) *The Relationship between Fertility Rate and Economic Growth in Developing Countries*(Master thesis). Retrieved from http://lup.lub.lu.se/luur/download?func=downloadFilerecordOId=8727479fileOId=8768892

[4] Gary S. Becker, Kevin M. Murphy, and Robert Tamura(1990) *Human Capital, Fertility, and Economic Growth* (Journal of Political Economy). Retrieved from https://www.journals.uchicago.edu/doi/citedby/10.1086/261723

[5] Douglas C. Montegomery, G. Geoffery Vining(2012). *Introduction to Linear Regression Analysis*. Hoboken, New Jersey, USA : John Wiley Sons, Inc.

[6] William H. Greene(2012). *Econometric Analysis*. New York, USA : PEARSON

Appendix A. Name of the countries

The 11 countries are as follows :

1. Antigua and Barbuda
2. Bermuda
3. Cuba
4. East Asia Pacific (excluding high income)
5. Gibraltar
6. Croatia
7. Kiribati
8. Maldives
9. Philippines
10. Palau
11. Pacific island small states

Appendix B. Regression Results

B.1. Initial Model Results

The results of the initial model,

$$Gdpchange = \beta_0 + \beta_1 GINI + \beta_2 gsavs + \beta_3 unemp + \beta_4 Fer + \epsilon$$

for 11 countries are as follows :

Country No.	Intercept	coeff. for unemp	coeff. for gsavs	coeff. for GINI	coeff. for fer
1	104.732771447	-1.187196945	0.211160663	0.008475085	-51.525583100
2	11.64271153	7.94837317	0.52660187	-0.05916092	-17.04734212
3	35.6069725	-0.1819797	0.6354548	-0.8749833	1.8241811
4	36.4247611	-1.3390286	-0.4483836	-0.4670971	2.9743344
5	-67.5668157	-2.2505394	0.0153976	3.4752971	-14.6989277
6	-24.6736997	0.8491430	0.3714272	0.4465316	-2.0833940
7	32.5916296	-1.2523247	-0.2039029	-0.1218005	-3.6826199
8	-114.2580237	2.0574644	0.9610662	-0.7022711	89.0429559
9	-69.2708809	-1.4615133	-0.1879790	1.7514666	-0.2497131
10	-24.2544429	3.2635599	1.1758308	-0.6487203	8.8510597
11	-27.1405674	-2.2086176	0.4301976	0.7820993	-1.0383089

Appendix C. Variance Covarinace Matrix for 11 Countries

The variance covariance matrix can be used to identify the correlation between regressors. The data for 11 countries is as follows :

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 18983.03049 -15.7372855 25.58902487 -80.54573389 -10047.520540
unemp[, i] -15.73729 0.4191075 0.17593746 0.32982125 -2.702117
gsavs[, i] 25.58902 0.1759375 0.18230743 0.01446736 -19.471410
gini[, i] -80.54573 0.3298212 0.01446736 0.74392579 31.139841
fer[, i] -10047.52054 -2.7021173 -19.47141034 31.13984059 5708.290499

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 714.684578 3.0365822 -2.67483592 -19.52061624 -70.2666108
unemp[, i] 3.036582 30.9122644 -0.18993894 -1.89030140 20.7332481
gsavs[, i] -2.674836 -0.1899389 0.04579298 0.06697369 -0.1699228
gini[, i] -19.520616 -1.8903014 0.06697369 0.67237898 0.3641882
fer[, i] -70.266611 20.7332481 -0.16992281 0.36418820 32.9366098

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 788.195244 -7.57308209 -7.04632288 -10.25772445 -63.053941
unemp[, i] -7.573082 0.26470944 0.18999906 -0.04792294 2.606818
gsavs[, i] -7.046323 0.18999906 0.31194636 -0.04980817 1.759530
gini[, i] -10.257724 -0.04792294 -0.04980817 0.28655848 -1.373843
fer[, i] -63.053941 2.60681755 1.75952985 -1.37384329 43.171803

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 651.701598 -3.33928495 -7.22199131 -3.77062505 -119.75320586
unemp[, i] -3.339285 1.35631383 0.03724491 -0.05264325 -1.14815644
gsavs[, i] -7.221991 0.03724491 0.17211507 0.07989994 -0.02039674
gini[, i] -3.770625 -0.05264325 0.07989994 0.81395113 -13.90860038
fer[, i] -119.753206 -1.14815644 -0.02039674 -13.90860038 302.91721430

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 1026.0391789 10.5955715 -0.81279414 -31.01980033 6.3241679
unemp[, i] 10.5955715 0.9677561 0.14097967 -0.73838256 0.8638643
gsavs[, i] -0.8127941 0.1409797 0.03722813 -0.05447088 0.1705681
gini[, i] -31.0198003 -0.7383826 -0.05447088 1.20250968 -1.7753349
fer[, i] 6.3241679 0.8638643 0.17056807 -1.77533491 25.0665351

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 64.02873338 -1.102629552 0.03915154 -1.395551383 5.32010222
unemp[, i] -1.10262955 0.203920131 0.02138758 0.001413988 -0.07505617
gsavs[, i] 0.03915154 0.021387575 0.01900777 -0.007044120 -0.03476416
gini[, i] -1.39555138 0.001413988 -0.00704412 0.038146343 -0.18416660
fer[, i] 5.32010222 -0.075056167 -0.03476416 -0.184166602 1.88001020

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 1655.233272 -41.05694159 -1.667813825 -17.697733078 -256.5735414
unemp[, i] -41.056942 1.53576870 0.034301566 0.329283336 6.0244758
gsavs[, i] -1.667814 0.03430157 0.048558854 0.009969469 0.1051907
gini[, i] -17.697733 0.32928334 0.009969469 0.250257174 2.4729982
fer[, i] -256.573541 6.02447582 0.105190719 2.472998161 44.5439763

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 10947.40181 -74.5841553 -43.9303150 31.8160137 -8552.51532
unemp[, i] -74.58416 2.1353996 0.6367226 -0.8982246 63.00951
gsavs[, i] -43.93032 0.6367226 0.3483757 -0.3308633 35.40022
gini[, i] 31.81601 -0.8982246 -0.3308633 0.5046414 -29.17850
fer[, i] -8552.51532 63.0095130 35.4002213 -29.1785012 6754.32318

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 5113.896498 8.00999505 -9.22967610 -2.5754338 -1835.425507
unemp[, i] 8.009995 0.41112705 0.08431691 -0.6748377 8.555374
gsavs[, i] -9.229676 0.08431691 0.06660188 -0.0546582 3.744952
gini[, i] -2.575434 -0.67483773 -0.05465820 2.0373187 -37.629086
fer[, i] -1835.425507 8.55537392 3.74495176 -37.6290862 1399.150656

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 619.518462 44.5577093 2.37351467 -1.6929013 -297.71740908
unemp[, i] 44.557709 8.4527355 0.95409551 -1.6409770 -7.99697454
gsavs[, i] 2.373515 0.9540955 0.15834467 -0.1998435 -0.01513367
gini[, i] -1.692901 -1.6409770 -0.19984355 0.5684943 -5.62945688
fer[, i] -297.717409 -7.9969745 -0.01513367 -5.6294569 231.70034785

(Intercept) unemp[, i] gsavs[, i] gini[, i] fer[, i]
(Intercept) 1454.390893 40.21828550 3.0513128 -33.27916987 -14.137953
unemp[, i] 40.218286 3.65225136 -0.8070090 -0.09509064 -15.208658
gsavs[, i] 3.051313 -0.80700901 0.6760678 -0.44602054 4.563141
gini[, i] -33.279170 -0.09509064 -0.4460205 1.27988700 -8.478879
fer[, i] -14.137953 -15.20865808 4.5631414 -8.47887878 162.415167

Appendix D. Subset Regression

D.1. Model excluding Unemployment

The VIF matrix of the model with the unemployment rate dropped from the initial data is as follows :

Country No.	VIF for GINI	VIF for gsavs	VIF for fer
1	1.704	2.241	2.648
2	1.644	1.233	1.444
3	1.214	1.015	1.213
4	6.345	1.306	6.043
5	1.579	1.459	1.677
6	2.725	1.458	2.564
7	1.596	1.004	1.596
8	1.309	2.008	1.631
9	6.199	3.383	5.621
10	15.879	1.351	16.774
11	5.933	2.309	4.016

D.2. Model excluding GINI

The VIF matrix of the model with the GINI coefficient dropped from the initial data is as follows :

Country No.	VIF for unemp	VIF for gsavs	VIF for fer
1	2.419	3.762	2.050
2	2.403	1.081	2.485
3	3.485	1.754	2.536
4	1.094	1.254	1.306
5	3.022	3.039	1.088
6	1.151	1.540	1.371
7	1.548	1.012	1.535
8	1.051	1.666	1.688
9	4.229	4.470	3.241
10	17.320	2.349	13.169
11	6.535	2.415	4.306

D.3. Model excluding Fertility Rate

The VIF matrix of the model with the fertility rate dropped from the initial data is as follows :

Country No.	VIF for unemp	VIF for gsavs	VIF for GINI
1	3.703	2.395	2.020
2	1.676	1.241	1.973
3	1.456	1.389	1.061
4	1.093	1.313	1.370
5	5.510	3.154	2.661
6	1.135	1.597	1.437
7	1.012	1.015	1.004
8	3.031	2.063	3.909
9	8.091	3.882	6.837
10	38.112	4.224	27.430
11	3.995	2.542	3.890

Appendix E. Results of Proposed Model

E.1. Proposed Model

$$Gdpchange = \beta_0 + \beta_1 fer + \beta_2 gsavs + \beta_3 unemp + \epsilon$$

The intercepts and the coefficients of the various regressors for the above model for 11 countries are as follows :

Country No.	Intercept	coeff. for unemp	coeff. for gsavs	coeff. for fer
1	105.65	-1.191	0.211	-51.880
2	9.250	7.782	0.534	-17.015
3	4.285	-0.328	0.484	-2.371
4	34.260	-1.369	-0.403	-5.010
5	22.081	-0.116	0.173	-9.568
6	-8.337	0.832	0.453	0.073
7	23.978	-1.092	-0.199	-2.480
8	-69.982	0.807	0.501	48.437
9	-67.057	-0.882	-0.141	32.099
10	-26.185	1.391	0.9477	2.427
11	-6.804	-2.151	0.710	4.143

Appendix F. Panel Data Regression

F.1. Pooled OLS Estimator

plm(formula = gdpchange unemp + gsavs + fer, data = pdata, model = ”pooling”)

Balanced Panel: n = 11, T = 15, N = 165

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-19.387579 -1.806098 0.097486 2.360309 9.580060

Coefficients:

Estimate Std. Error t-value Pr(>—t—)

(Intercept) 1.757786 2.055571 0.8551 0.393748

unemp 0.021135 0.077711 0.2720 0.785998

gsavs 0.197493 0.060496 3.2646 0.001339 **

fer -0.278788 0.497218 -0.5607 0.575785

—

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares: 2684.2

Residual Sum of Squares: 2485.4

R-Squared: 0.074067

Adj. R-Squared: 0.056813

F-statistic: 4.29288 on 3 and 161 DF, p-value: 0.0060531

F.2. Between Estimator

Oneway (individual) effect Between Model

Call:

plm(formula = gdpchange gsavs + unemp + fer + year$_i dx, data = pdata, model =$ ”$between$”)

Balanced Panel: n = 11, T = 15, N = 165

Observations used in estimation: 11

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-0.617159 -0.364343 0.093914 0.269437 0.640440

Coefficients:
Estimate Std. Error t-value Pr(>—t—)
(Intercept) -0.616012 1.393318 -0.4421 0.671741
gsavs 0.209882 0.047027 4.4630 0.002926 **
unemp 0.183070 0.043488 4.2097 0.003987 **
fer 0.152849 0.264743 0.5773 0.581789
—
Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Total Sum of Squares: 8.5508
Residual Sum of Squares: 1.727
R-Squared: 0.79803
Adj. R-Squared: 0.71147
F-statistic: 9.21933 on 3 and 7 DF, p-value: 0.0079131

F.3. Within of Fixed Effects Estimator

Call:
plm(formula = gdpchange   unemp + gsavs + fer, data = pdata, model = "within")
Balanced Panel: n = 11, T = 15, N = 165
Residuals:
Min. 1st Qu. Median 3rd Qu. Max.
-17.5603 -2.0581 0.3480 2.3440 8.4059

Coefficients:
Estimate Std. Error t-value Pr(>—t—)
unemp -0.612706 0.170304 -3.5977 0.0004345 ***
gsavs 0.103364 0.080022 1.2917 0.1984396
fer -1.732211 1.488036 -1.1641 0.2462225
—
Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Total Sum of Squares: 2555.9
Residual Sum of Squares: 2162.9
R-Squared: 0.15376
Adj. R-Squared: 0.080906

F-statistic: 9.14553 on 3 and 151 DF, p-value: 1.3415e-05

F.4. Random Effects Estimator

Call: plm(formula = gdpchange unemp + gsavs + fer, data = pdata, model = "random")

Balanced Panel: n = 11, T = 15, N = 165

Effects:

var std.dev share

idiosyncratic 14.324 3.785 1

individual 0.000 0.000 0

theta: 0

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-19.387579 -1.806098 0.097486 2.360309 9.580060

Coefficients:

Estimate Std. Error t-value Pr(>—t—)

(Intercept) 1.757786 2.055571 0.8551 0.393748

unemp 0.021135 0.077711 0.2720 0.785998

gsavs 0.197493 0.060496 3.2646 0.001339 **

fer -0.278788 0.497218 -0.5607 0.575785

—

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Total Sum of Squares: 2684.2

Residual Sum of Squares: 2485.4

R-Squared: 0.074067

Adj. R-Squared: 0.056813

F-statistic: 4.29288 on 3 and 161 DF, p-value: 0.0060531