**Table of Contents**

**1. Introduction**

**2. Literature Review**

2.1. Dependent Variable: Gross Domestic Savings (% of GDP)

2.2. Independent Variables

  a. Impact of GNI per Capita on Saving Rate

  b. Impact of Age Dependency Ratio on Saving Rate

  c. Impact of Inflation on Saving Rate

  d. Impact of Real Interest Rate on Saving Rate

**3. Methodology**

3.1. Model Specification

 3.2. Data Description

**4. Ordinary Least Squares (OLS) Regression Analysis**

 4.1. Assumption 1: Linearity

 4.2. Assumption 2: No Perfect Multicollinearity

 4.3. Assumption 3: Zero Conditional Mean

 4.4. Assumption 4: Homoscedasticity

  4.4.1. Scatter Plot of Residuals

  4.4.2. Breusch–Pagan Test

  4.4.3. White’s Test

  4.4.4. Robust Test

 4.5. Assumption 5: No Autocorrelation

 4.6. Assumption 6: Normality of Error Terms

**5. Stationarity and Cointegration Analysis**

5.1. Augmented Dickey-Fuller (ADF) Unit Root Test

 5.2. Johansen Cointegration Test

 5.3. Vector Error Correction Model (VECM)

  5.3.1. Long-Run Effects

  5.3.2. Short-Run Effects

**6. Conclusion and Policy Implications**

6.1. Conclusion

 6.2. Policy Implications

**7. Limitations and Future Research Recommendations**

**References**

**Appendices**

# **ABSTRACT**

The paper examines the factors affecting the gross domestic saving rate in the United States between 1975 and 2021 such as the influence of real interest rates, inflation, the ratio of age dependency and the level of income per capita. Based on time-series data, the paper will verify that there exist long-run relationships between the significant variables with the assistance of Ordinary Least Squares (OLS) and Vector Error Correction Model (VECM). The findings indicate that the age dependency ratio and real interest rates have negative and significant consequences on saving rate in the long run and the per capita income and inflation have no significant effects. They have limited impacts in the short-run, but because the error correction term is substantial, this implies adjustment to the equilibrium gradually. These results underline the influence of the structural demographic and financial factors in constructing the national saving behavior. The policy implications are that demographic pressure needs to be taken care of and there should be macro-financial stability in order to promote long-term domestic saving.

# **1. INTRODUCTION**

There are a number of studies that examined the US domestic saving rate (% of GDP). Bovenberg & Evans (1990) realized that the real interest rates and demographics determined saving in the 1980s. According to the IMF (Ouliaris & Rochon, 2018) there were no significant structural developments after the 2008 crisis and the rise in saving activity during the period of 2008-2012 was primarily based on the income shocks and wealth. Since the process of saving allocation to investment, financial stability, and long-term development is widely discussed, it is essential to conduct further studies during other time frames like 1975-2021 in this study.

# **2. LITERATURE REVIEW**

**2.1. Dependent Variable:**

**Gross domestic savings (% of GDP):**

Gross domestic savings reflects the portion of GDP remaining after spending and the ability to invest in the long term (Glossary | DataBank, 2015). In the US, the period 1970–2022 ranged from 13.9% to 23.8%, averaging 19–20% (TheGlobalEconomy.com). The period 1975–2021 was volatile, suitable for analyzing the relationship with macroeconomic factors.

**2.2. Independent Variables:**

**a, Impact of GNI on Savings rate**

GNI per capita is national income (including GDP and net income from abroad) divided by the population, reflecting the average standard of living. Loayza et al. (2000) supports that GNI is expected to exert a positive influence on saving, as higher income levels enhance individuals’ capacity to generate surplus. However, Brueckner et al. (2021)’s research on 130 countries over the period 1960–2017 shows that in rich countries, increasing GNI can slow down saving. Therefore, in the US, a high-income country, this relationship can be negative or nonlinear.

**b, Impact of Age dependency ratio on Savings rate**

The age dependency ratio is the number of people under 15 and over 65 years old compared to the working age group (15–64). The age dependency ratio is anticipated to have a negative effect on savings, since a larger dependent population reduces the share of income that can be allocated to saving, leading to increasing financial burden (Hyung, 2013). Moreover, another study also asserts that a high dependency ratio negatively affects savings, especially when the dependent group exceeds the working age group (Li et al., 2006).

**c, Impact of Inflation, consumer prices on Savings rate**

Inflation (CPI) is the annual percentage change in the cost of a standard basket of consumer goods and services. Davidson & MacKinnon (1983) argue when inflation is high, nominal interest rates include an inflation premium, causing the saving index to increase, even though the actual accumulated assets do not increase.

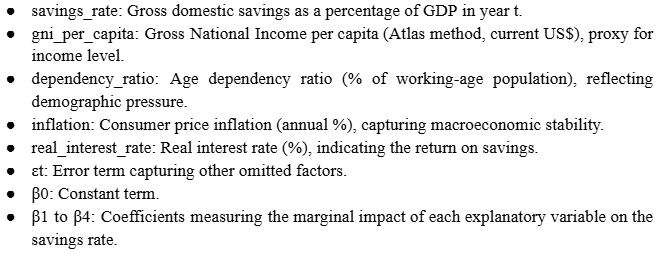
**d, Impact of Real interest rate on Savings rate**

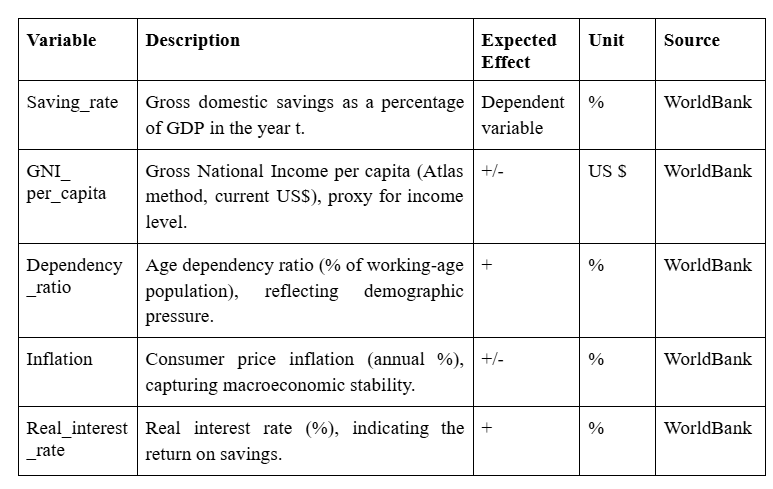
Real interest rates reflect the actual returns that savers receive after deducting inflation, and play a role in regulating saving and consumption behavior. Therefore, real interest rates are expected to encourage saving by increasing the return to deferred consumption, in line with intertemporal substitution theory. Research by Khan, Khilji & Abbas (2014) shows that real interest rates have a positive impact on the savings rate: when real interest rates increase, the opportunity cost of current consumption increases, promoting savings and contributing to national financial stability.

# **3. METHODOLOGY**

**3.1. Model specification (Appendix 1)**

This study employs a multiple linear regression model to examine the determinants of gross domestic savings, using the Ordinary Least Squares (OLS) estimation method. The dependent variable is the gross domestic savings rate (% of GDP), while the explanatory variables include gross national income (GNI) per capita, the age dependency ratio, consumer price inflation, and the real interest rate. The model is specified as follows:



This model is theoretically grounded in the life-cycle hypothesis (Modigliani and Brumberg, 1954) and the permanent income hypothesis (Friedman, 1957), which suggest that saving behavior reflects efforts to smooth consumption across the life course based on anticipated long-term income. 

**3.2. Data description:**

This study employs annual time-series data for the United States spanning from 1975 to 2021, resulting in a total of 47 observations. All data are obtained from the World Development Indicators, ensuring consistency and cross-country comparability.

Descriptive statistics for the main variables are presented in Table 4.2 (Appendix 2.1). Over the sample period, the gross domestic savings rate in the U.S. averaged 19.55% of GDP. This reflects meaningful fluctuations in saving behavior that likely respond to macroeconomic, policy, and demographic dynamics.

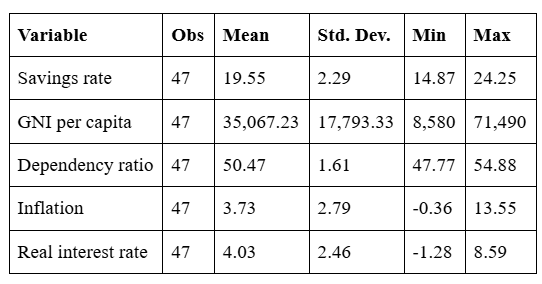
GNI per capita, measured using the Atlas method in U.S. dollars, had an average of 35,067 USD. The standard deviation is notably high (17,793 USD), reflecting the long-term income growth in the U.S. economy. The minimum and maximum values—8,580 USD and 71,490 USD—highlight the significant expansion of income over nearly three decades.

The age dependency ratio remained stable, averaging 50.47% with a narrow range, indicating a steady demographic structure throughout the sample. In contrast, inflation shows notable volatility, ranging from -0.36% (deflation) to 13.55%. This variability underscores the presence of both inflationary pressures and macroeconomic stabilization over time.

The real interest rate, a key determinant of saving incentives, ranged from -1.28% to 8.59%. The broad dispersion reflects shifts in monetary policy regimes, inflation expectations, and global financial conditions that affected the real return on savings.

Overall, the dataset offers sufficient variability across key macroeconomic indicators to support robust econometric analysis of their impact on saving behavior in the U.S. context.

**Table 4.2**



# **4. ORDINARY LEAST SQUARES (OLS) REGRESSION**

**4.1. Assumption 1: Linearity**

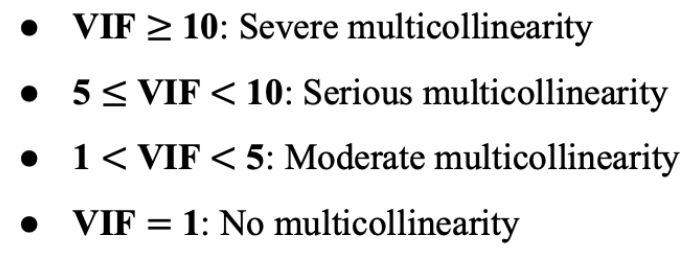


The first assumption is linearity in parameters, which means that the dependent variable and the explanatory variables should be linearly related in the coefficients (Wooldridge, 2017). We use Scatter Plots with Linear Regression Line as a tool to visualise the relationship between the dependent variable and the explanatory variables.

The STATA output (Appendix 3) demonstrates that the correlation between Savings\_rate and Gni\_per\_capita is negative, showing that two variables are indirectly related, whereas Dependency\_ratio, Inflation, and Real\_interest\_rate have rather stable positive slopes, indicating a weaker positive association. These scatter with fitted values validate the linearity assumption in the model, indicating that relationships between the explained variable and predictor variables are correct.

**4.2. Assumption 2: No Perfect Collinearity**

Multicollinearity is a condition where independent variables are closely connected and may skew the reliability of the estimations of coefficients. The Variance Inflation Factor (VIF) is a statistical measure applied to identify such correlation. According to Wooldridge (2017):

****

**Null hypothesis (H0):** There is no perfect multicollinearity

**Alternative hypothesis (H1):** There is perfect multicollinearity

Since all variables have VIF values less than 5 and the mean VIF is only 1.69, this demonstrates a low multicollinearity in the model (Appendix 4). According to Gujarati and Porter (2009), if perfect collinearity were present, at least one VIF value would become undefined (as R² equal to 1). Therefore, the model does not have perfect multicollinearity, which supports the reliability of the estimated coefficients.

**4.3. Assumption 3: Zero Conditional Mean**

The Ramsey RESET is used to determine whether there are any omitted variables or misspecifications, especially the non-linearities that are not captured by the fitted values of the dependent variable.

**Null hypothesis (H0):** E(∣​) = 0 (Regressors are exogenous)

**Alternative hypothesis (H1):** E(∣​) 0 (Regressors are endogenous)

The STATA result (Appendix 5) indicated an F-statistic of 1.43 and a p-value of 0.2479. Since the p-value is greater than the significance level of 5%, we can not reject the null hypothesis so there are no omitted variables in the model. Therefore, the OLS regression model is well specified and does not need the consideration of the non-linear terms or the higher powers of the fitted values.

**4.4. Assumption 4: Homoscedasticity**

**Null hypothesis (H0):** There is homoskedasticity

**Alternative hypothesis (H1):** There is heteroskedasticity

**4.4.1. Scatter Plots**

The heteroscedasticity has an impact on the regression model, so both visual inspection checks and formal statistical testing were used to verify the correctness of the results. A scatter plot of residuals against the fitted values was created to conduct the visual inspection (Appendix 6.1). The plot has the residuals randomly spread with no noticeable trends, indicating that the variance of the residuals is identical at different levels of the fitted values. This shows that there is no heteroscedasticity.

**4.4.2. Breusch-Pagan Test**

The Breusch-Pagan test is a method used to indicate whether the variance of the residuals is constant across observations. This test is applied to estimate a null hypothesis that the residuals are homoskedastic (Wooldridge, 2017). The STATA result indicates a chi-square statistic value of 1.24 and a p-value of 0.2663 (Appendix 6.2), thus, the null hypothesis is accepted stating that there is no heteroskedasticity in the model.

**4.4.3. White’s Test**

In addition, we applied White’s test, which is stronger against functional form misspecification. The null hypothesis of homoskedasticity is also tested by White's test. The outcome of a p-value of 0.0360 and a chi-square value of 31.45 (Appendix 6.3). The null hypothesis is rejected at the 5% level of significance since the p-value is less than 0.05, suggesting that the model may contain heteroskedasticity.

**4.4.4. Robust test**

The heteroskedasticity can lead to inefficient OLS estimates and invalid measurement of the standard errors, in which the inference is inaccurate. The solution to this problem is working with the robust test, which can be used to provide valid inference in cases where the error variance is not constant (Wooldridge, 2016). From Appendix 6.4, the test obtains a p-value of 0.000 and an F-statistic of 56.36, which shows that the model is mutually significant at a 1% level. The robust estimation improves the strength of the outcome of the F-test, supporting the overall explanatory ability of the model.

**4.5. Assumption 5: No Autocorrelation**

**Null hypothesis (H0):** Cov (​,​) = 0 (There is no first-order autocorrelation)

**Alternative hypothesis (H1):** Cov (​,​) 0 (There is first-order autocorrelation)

To evaluate if this model violates assumption 5, we divide this into 2 steps. Firstly, we conducted a “Lag-order selection test” to choose the appropriate optimal lag length since autocorrelation is strongly sensitive to the model specification, especially for the time series data used in this case. From the STATA’s result, we selected lag 2 as an optimal lag length because this lag generates the lowest value of 3 criteria (AIC,BIC, and HQIC) (Appendix 7.1). After identifying the appropriate lag length, two lagged values of the dependent variable were included in the model and the Durbin-Watson test is run to detect autocorrelation. A d-statistic of 1.933314, which is extremely close to 2, was gathered from the test, indicating that the regression model's residuals do not show substantial autocorrelation, and we accept the null hypothesis.(Appendix 7.1)

**4.6. Assumption 6: Normality**

**Null hypothesis (H0):** The error terms are normally distributed

**Alternative hypothesis (H1):** The error terms are not normally distributed

The Skewness-Kurtosis test will determine whether residuals are normally distributed, which measures two parameters: skewness - the presence of asymmetry, and kurtosis - measure of the thickness of the tails. The p-value of skewness = 0.7198 and kurtosis = 0.6453 based on the normality test (Appendix 8), which constitutes 47 observations. This is similarly the case with the joint test statistic of a p-value of 0.8434. The null hypothesis of normality cannot be rejected because all of the p-values are higher than the significance level of 5%. This result reflects on the model inference reliability since there are no significant deviations of residuals to be considered non-normal. It is also essential to check the size of the normality because most of the tests that use regression require normally distributed errors to make a valid statistical inference (Wooldridge, 2016; Gujarati & Porter, 2009). Since the normality test results indicate that the residuals of the model do not show any significant deviations, it follows that the regression model assumption seems to be satisfactory.

**5. Stationary Test:**

**5.1 Stationary Test (Appendix 9)**

To examine the time-series properties of its variables, ADF tests of all the variables at level were run, with two lag and a deterministic trend. The findings show that all variables, the Saving\_rate, GNI\_per\_capita, Dependency\_ratio, Inflation and Real\_interest\_rate are not stationary in levels 0. In particular, all the ADF tests statistics are greater than the 5% critical values and the MacKinnon approximate p-values are above 0.05. As an example, the test statistic in the case of Saving\_rate is -2.505 against 5 percent critical value which is -3.524 (p = 0.3252). In the same way, Dependency ratio and GNI\_per\_capita are 0.4073 and 0.9837 respectively.

The results validate the existence of a unit root in all the variables which means the variables are integrated of order 1 or I(1). This gives reason to follow it up with the Johansen cointegration test and estimation of a Vector Error Correction Model (VECM) which are suitable in operations when all of the series are non-stationary at the level, but might be cointegrated.

**5.2 The Johansen test (Appendix 10)**

The Johansen cointegration test was done to study the presence of long-run equilibrium conditions in the relationships between the variables. The null hypothesis was the absence of cointegring relationships (rank = 0) which was rejected in the light of the trace statistics and the critical values which were at 5%, in favor of rank <= 2. Nevertheless, the null hypothesis at rank = 2 was not rejected (Trace = 23.8946 < 29.68) and this shows two statistically significant cointegrating vectors. This observation supports the usage of a Vector Error Correction Model (VECM) consisting of rank = 2 to characterize both the dynamics in the short-run and equilibrium links in the long-run.

**5.3 VECM model (Appendix 11)**

**5.3.1 Long-run effect**

The cointegration findings reveal that the Dependency\_ratio and Real\_interest\_rate have significant negative effects on the national saving rate in long run and they are all statistically significant. In particular, the Saving\_rate decreases by about 1.1013 units as the Dependency\_ratio increases by one unit, and Saving\_rate decreases by about 0.6785 units when the Real\_interest\_rate is increased by one percentage point. These results are as expected theoretically: higher dependency burden has a tendency of lowering saving capacity because people need to consume more, and higher real interest rate may deter saving when there is the risk of being related to economic insecurity or stricter lending terms. Conversely, Inflation presents a negative coefficient, which was not statistically significant, and GNIpercapita does not even appear in the cointegrating equation of Savingrate, which lets one conclude about the absence of a long-run relationship.

**5.3.2. Short run effect**

In the short run, no effects of any of the lagged explanatory variables GNI per capita, Dependency ratio, Inflation, and Real interest rate have any scintilla of significance on the first difference of Saving rate. This implies that variation of these variables do not have any instant effect on the change in the saving rate. Nevertheless, the estimated error correction term of the second cointegrating equation (\_ce2) is significant and negative and, thus, indicates that the removal of disequilibrium of all of the cointegrating -relations is partially reflected in Saving\_rate. The adjustment mechanism here denotes that in spite of the fact that individual short run effects are small, there is a long run response that occurs in Saving\_rate to again reach long run equilibrium. This supports the significance of the cointegrated system and underlines the role of Saving\_rate as an adaptive variable in the macroeconomic system in general.

**6. Conclusion and Policy implication**

**6.1. Conclusion**

The report is an empirical study that investigated factors that explain gross domestic saving rate in the U.S. during the period of 1975 to 2021. The integration of macroeconomic and age dependency ratio, real interest rates, inflation, and per capita income as the determinants of saving behavior validate the implications that the saving behavior is affected by both structural and policy based factors.

The OLS regression and the stationarity tests as well as Johansen cointegration from the stationarity tests indicate a suitable use of a Vector Error Correction Model (VECM). The long run results show that the rate of dependency on aged population and real interest rates have significant negative impacts on the rate of national saving. The revenue per capita and inflation do not have long-term statistically significant impact to their surprise as well. There is no evidence of explosions in savings due to the short run effects of the explanatory variables; however, there is a significant error correction term suggesting that there is indeed some force of disequilibrium adjustments. This evidence demonstrates the importance of long-term structural condition instead of short-term shocks in the national saving behavior.

**6.2. Policy Implication**

The results indicate that demographic set up is one of the main factors that affect country savings. The policies should be introduced to cope with the demands of aging by reforming pension system and providing incentives to take part in the workforce. Implication of the negative impact of real interest rates sheds more light on the need to maintain financial stability and people confidence in monetary policy. Although the growth in income might not necessarily increase savings, the goal of financial security and discouraging income fluctuations can be used to foster the long-term effects of saving behavior.

**7. Limitation and Recommendations**

Although the present study is of significance, a few limitations have to be noted. On the one hand, the dataset provides a long history between 1975 and 2021 including some significant macroeconomic shocks into the picture like the 2008 global financial crisis or the COVID-19 pandemic. Nevertheless, the structural breaks or the regime shifts induced by these shocks are not strictly considered within the current model that can influence the stability of the relationships estimated. Second, the research is limited to macro-level data in one nation and is thus restricted regarding the external validity of the results. Third, it is assumed that only linear relationships exist, which can be too simplistic in explaining complicated dynamics in the behavior of saving. In addition, money distribution, household debt, and financial market access which may be significant factors are unfortunately not available in the data.

Potential directions of future research would be to apply structural break tests or dummy variables to control well-established economic crises. More refined dynamics potentially can be characterized by nonlinear models, including threshold regressions or time-varying parameter models. The comparative examination based on various economies or the panel data method would also enhance the study and provide more practical conclusions.

# **REFERENCES**:

BOVENBERG, A. L., & EVANS , O. (1990). *National and Personal Saving : Measurement and Analysis of Recent Trends* . International Monetary Fund. <https://www.elibrary.imf.org/display/book/9781557752314/ch002.xml>

Brueckner, M., Kikuchi, T., & Vachadze, G. (2021). Transitional dynamics of the saving rate and economic growth. *Macroeconomic Dynamics*, 1–24. <https://doi.org/10.1017/s1365100521000493>

Davidson, R., & MacKinnon, J. G. (1983). Inflation and the savings rate. *Applied Economics*, *15*(6), 731–743. <https://doi.org/10.1080/00036848300000062>

Friedman, M. (1957). *A theory of the consumption function*. Princeton University Press. Retrieved from<https://econpapers.repec.org/RePEc:nbr:nberbk:frie57-1>

Globalen LLC. (2022). *USA Savings, percent of GDP - data, chart | TheGlobalEconomy.com*. TheGlobalEconomy.com. <https://www.theglobaleconomy.com/USA/savings/>

*Glossary | DataBank*. (2015). Worldbank.org. <https://databank.worldbank.org/metadataglossary/world-development-indicators/series/NY.GDS.TOTL.ZS?utm>

*Gross national income (GNI) per capita*. (2021). Our World in Data. https://ourworldindata.org/grapher/gross-national-income-per-capita-worldbank

Hyung, J. (2013). *An Analysis on the Effect of Old Age Dependency Ratio on Domestic Saving Rate*. <https://econ.berkeley.edu/sites/default/files/Jinwoo%20Hyung.pdf>

Khan, I. H., Ahmad, & Abbas, G. (2014). RELATIONSHIP BETWEEN INTEREST RATE AND HOUSEHOLD’S SAVINGS IN PAKISTAN. *International Journal of Innovation and Scientific Research*, *12*(1), 295–308. <https://ijisr.issr-journals.org/abstract.php?article=IJISR-14-291-02>

Kremers, J. J. M., Ericsson, N. R., & Dolado, J. J. (1992). THE POWER OF COINTEGRATION TESTS. *Oxford Bulletin of Economics and Statistics*, *54*(3), 325–348. <https://doi.org/10.1111/j.1468-0084.1992.tb00005.x>

Leybourne, S. J. (1995). *TESTING FOR UNIT ROOTS USING FORWARD AND REVERSE DICKEY-FULLER REGRESSIONS.* Ebsco.com; Oxford Bulletin of Economics & Statistic. 10.1111/j.1468-0084.1995.tb00040.x

Li, H., Zhang, J., & Zhang, J. (2006). Effects of longevity and dependency rates on saving and growth: Evidence from a panel of cross countries. *Journal of Development Economics*, *84*(1), 138–154. <https://doi.org/10.1016/j.jdeveco.2006.10.002>

Loayza, N., Schmidt‑Hebbel, K., & Servén, L. (2000). *What drives private saving across the world?* The Review of Economics and Statistics, 82(2), 165–181. Retrieved from <https://papers.ssrn.com/abstract=629158>

Ouliaris, M. S., & Rochon, M. C. (2018). *The U.S. Personal Saving Rate*. Google Books. <https://books.google.com.vn/books?hl=vi&lr=&id=tbMZEAAAQBAJ&oi=fnd&pg=PA4&dq=The+U.S.+Personal+Saving+Rate1&ots=b_vbTs4L_h&sig=ki12kP6xneWUMPLidjM0YDMN5RE&redir_esc=y#v=onepage&q=The%20U.S.%20Personal%20Saving%20Rate1&f=false>

Prabhakaran, S. (2019, November 2). *Augmented Dickey Fuller Test (ADF Test)*. Machine Learning Plus. <https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/>

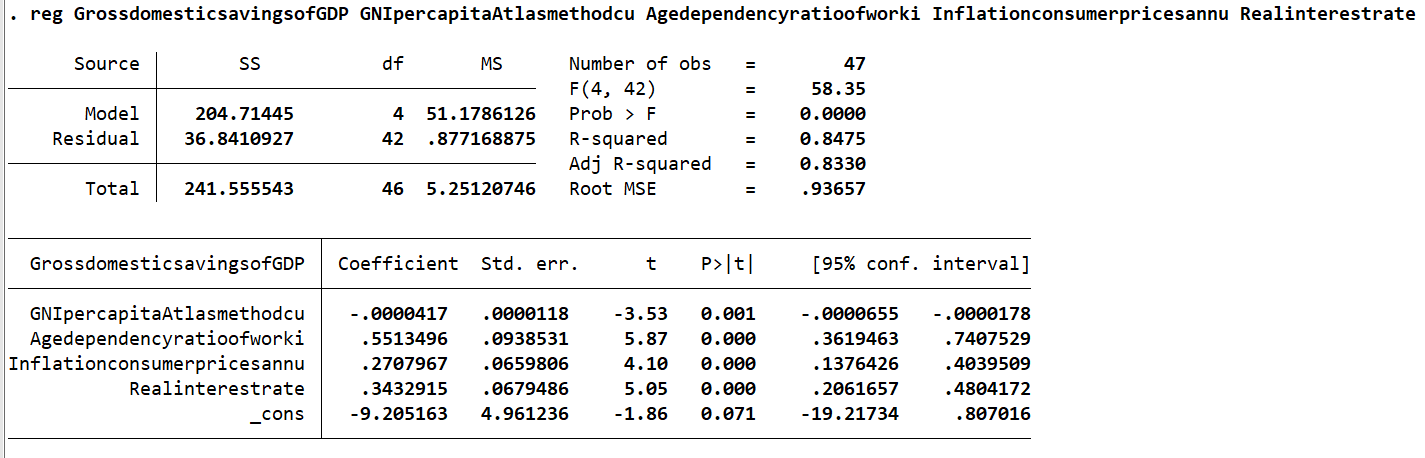
Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross‑section data. In K. K. Kurihara (Ed.), *Post‑Keynesian economics* (pp. 388–436). Rutgers University Press. Retrieved from<https://www.jstor.org/stable/1817129>

Salman, A., & Zaib, R. (2012). Age dependency and its effect on national savings: Evidence from selected Asian countries. *International Journal of Business and Management, 7*(12), 152–162. <https://ccsenet.org/journal/index.php/ijbm/article/view/12654>

World Bank. (1995). *World development report 1995: Workers in an integrating world*. Oxford University Press. Retrieved from <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/739011468169743467>

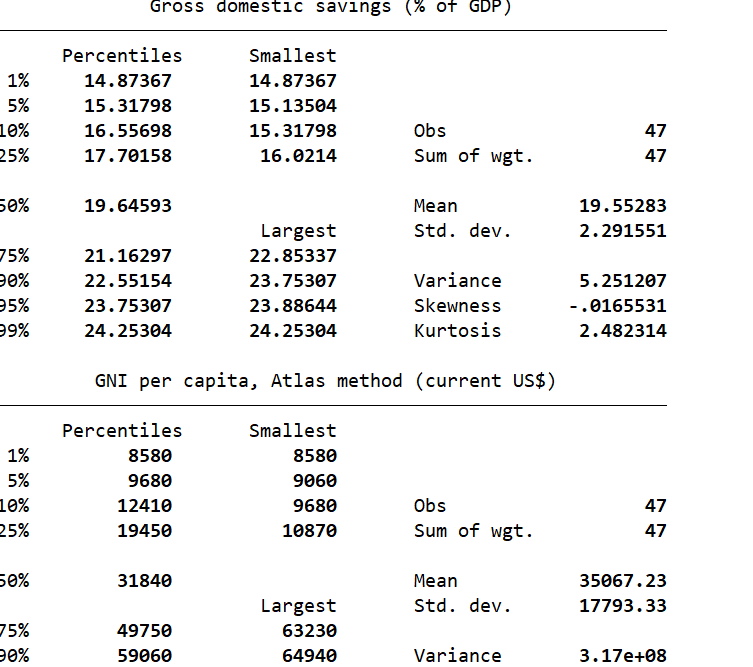
# **APPENDICES**:

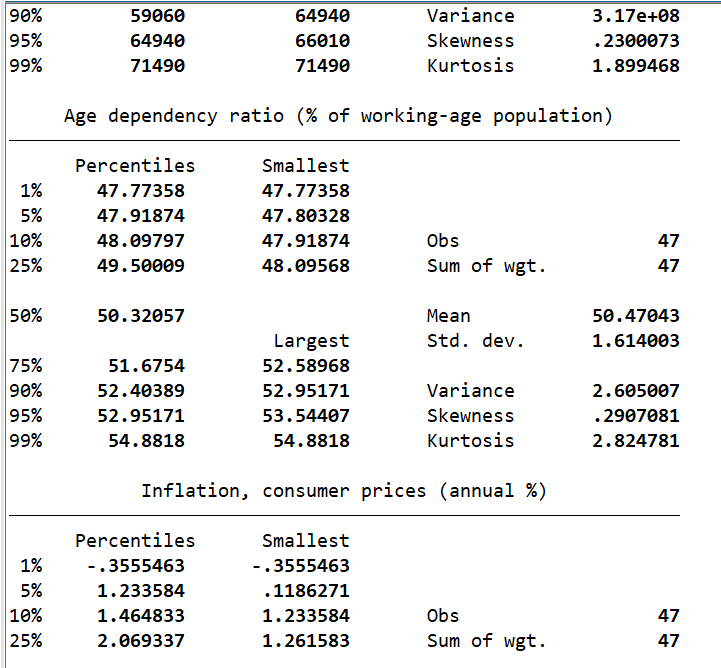
## **Appendix 1: Econometric Model**

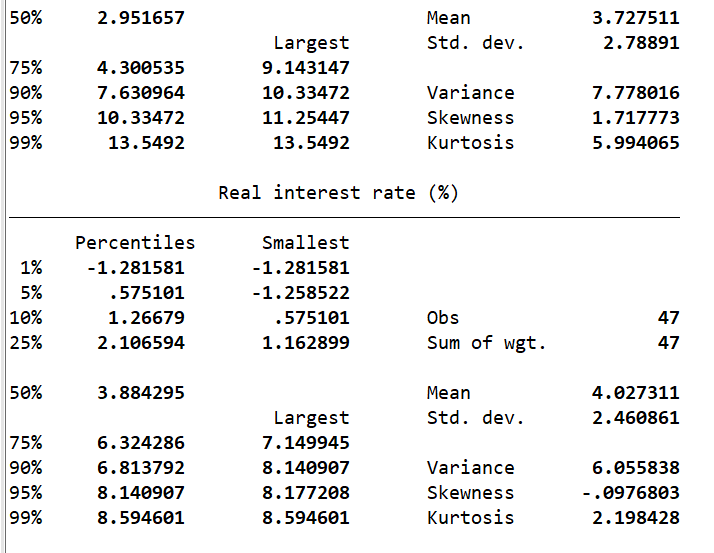
****

## **Appendix 2:**

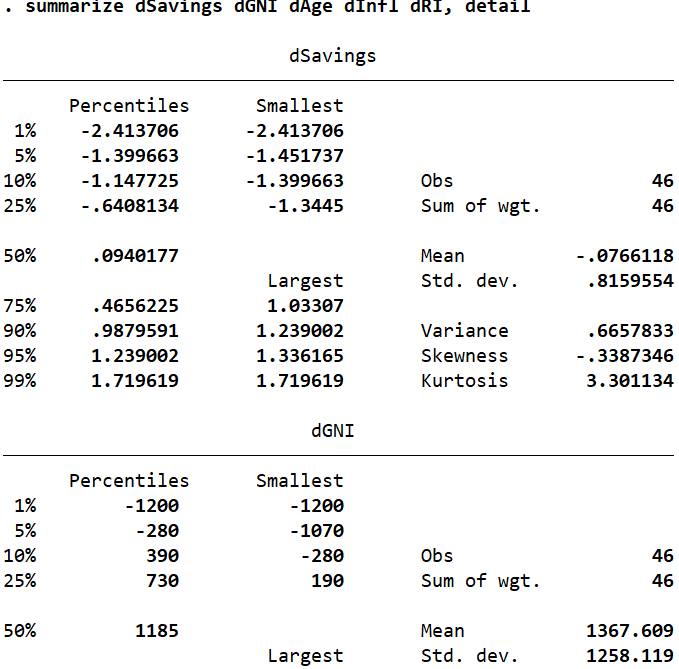
**Appendix 2.1:**

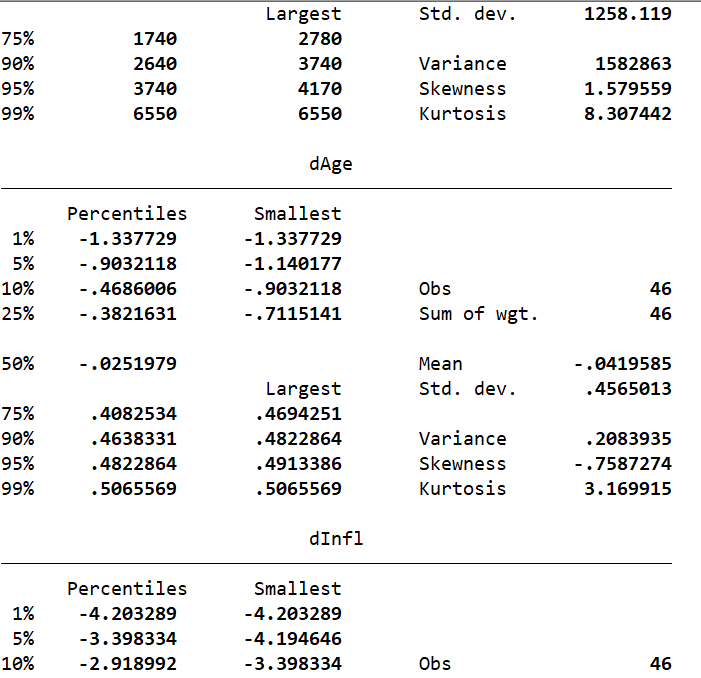
****

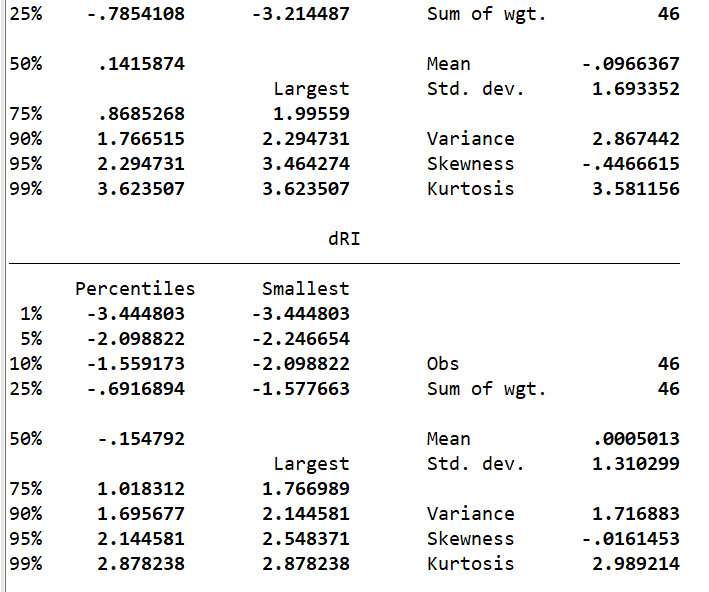
****

****

**Appendix 2.2:**

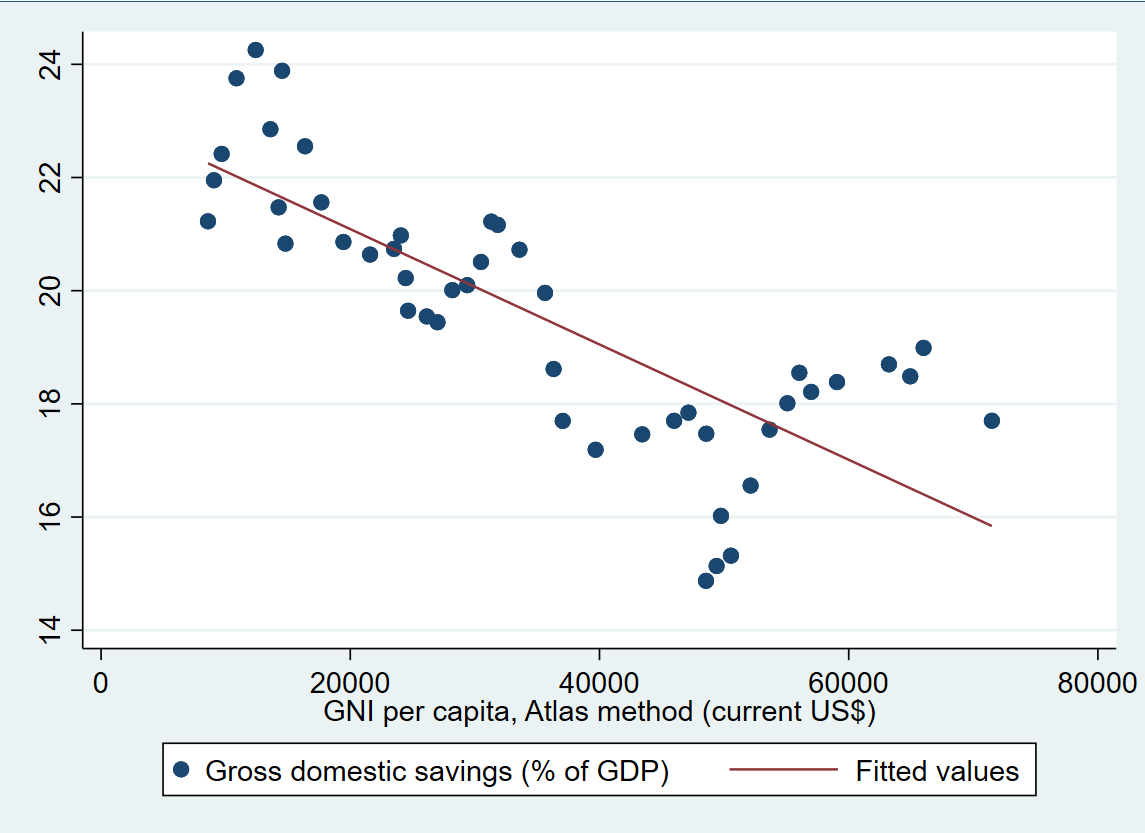
****

****

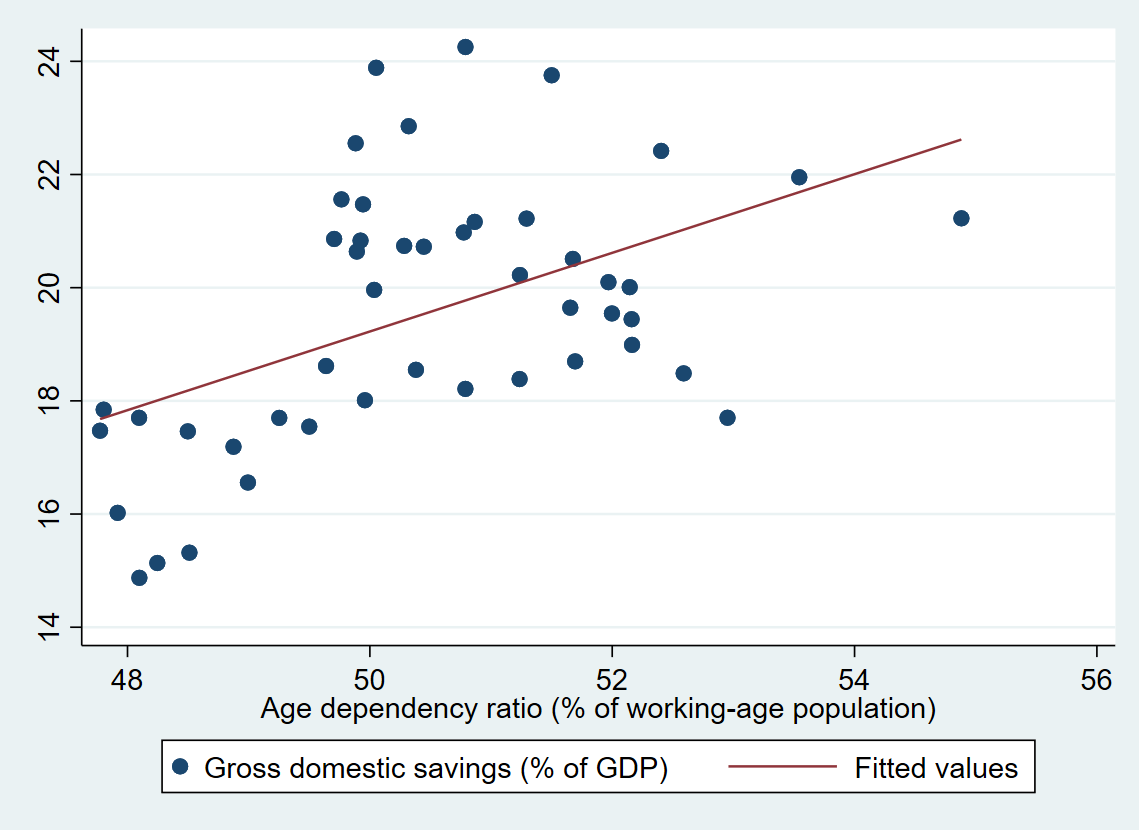
****

## **Appendix 3: Linearity**

**Appendix 3.1: Scatter and Fit for Savings and GNI**

****

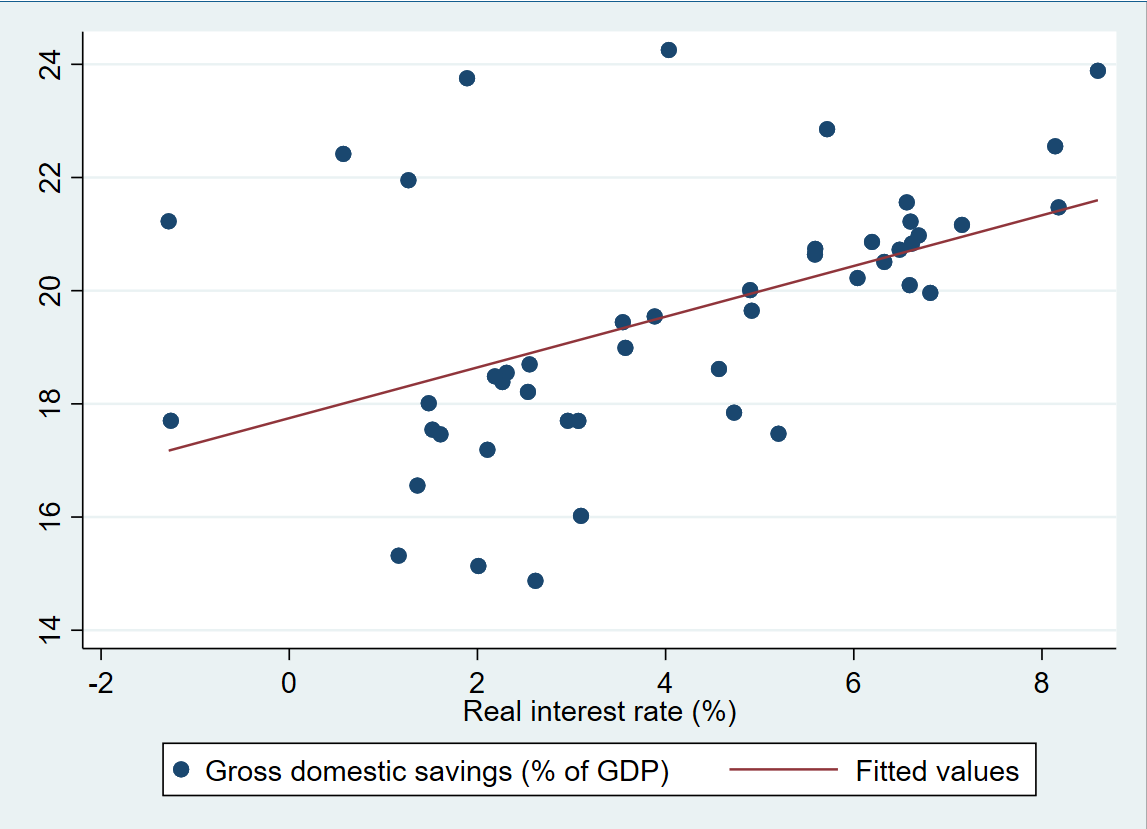
**Appendix 3.2: Scatter and Fit for Savings and Age**

****

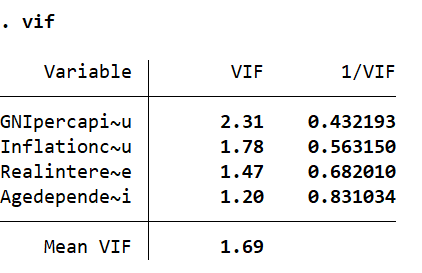
**Appendix 3.3: Scatter and Fit for Savings and Inflation**

****

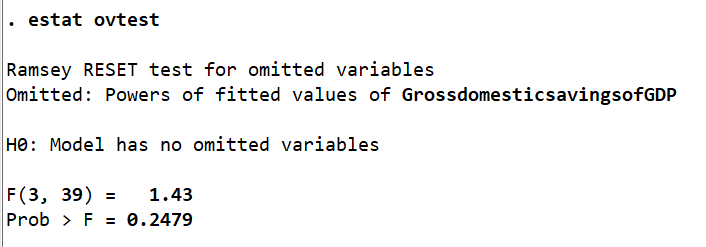
**Appendix 3.4: Scatter and Fit for Savings and Rate Interest**

****

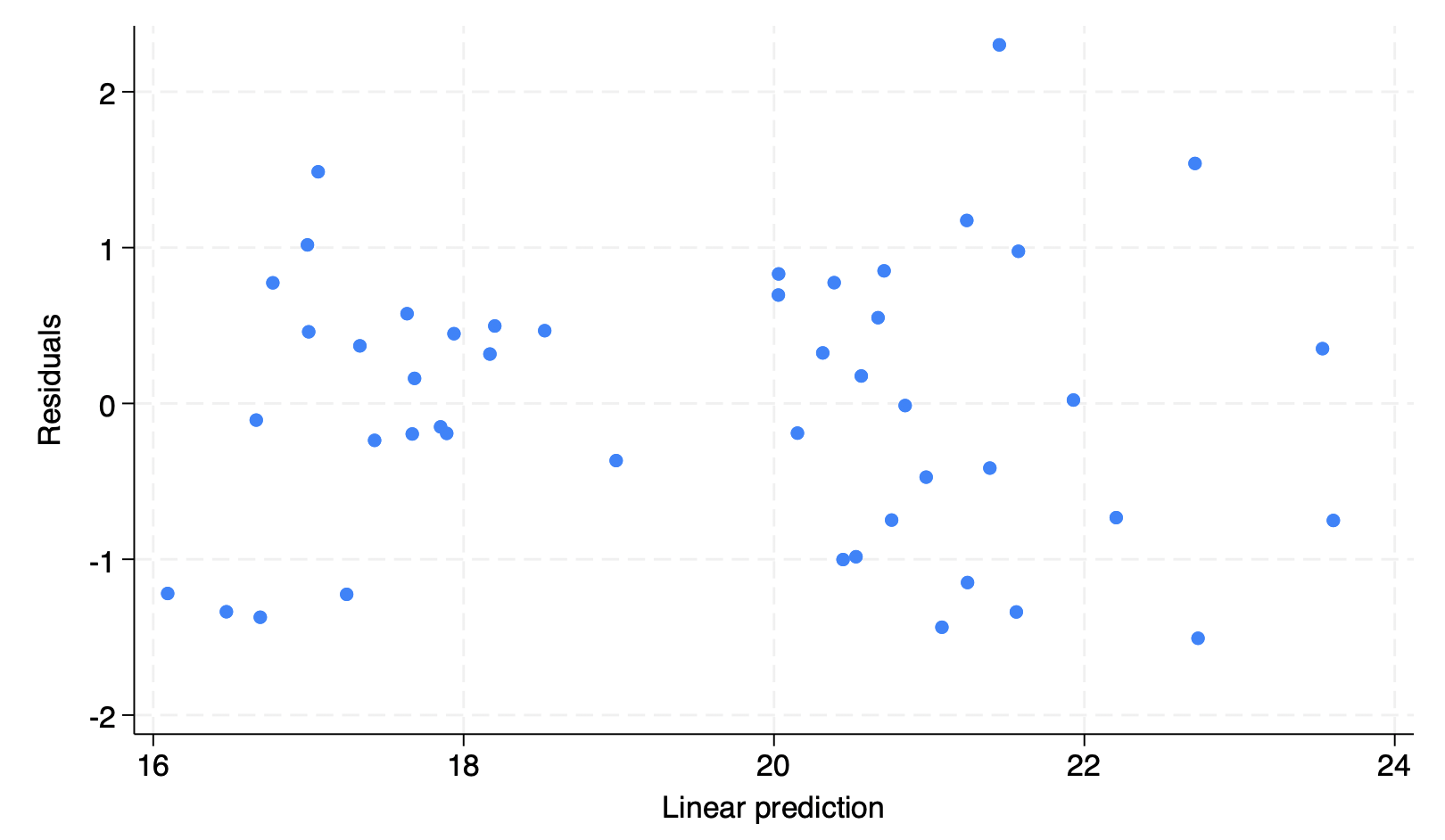
## **Appendix 4: VIF**

****

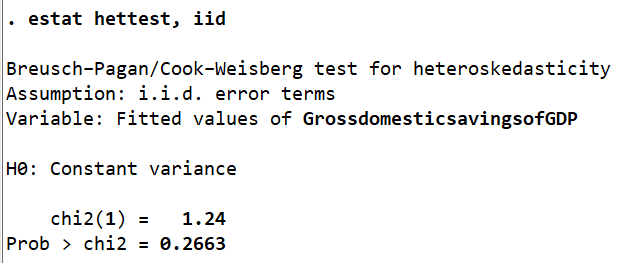
**Appendix 5: Ramsay RESET**

****

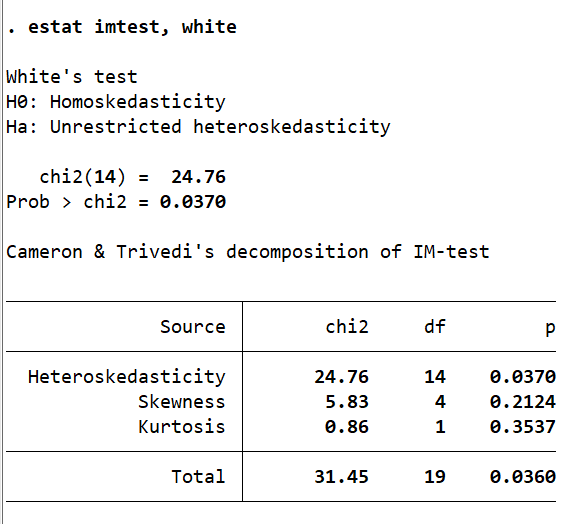
**Appendix 6.1: Scatter Plots**

****

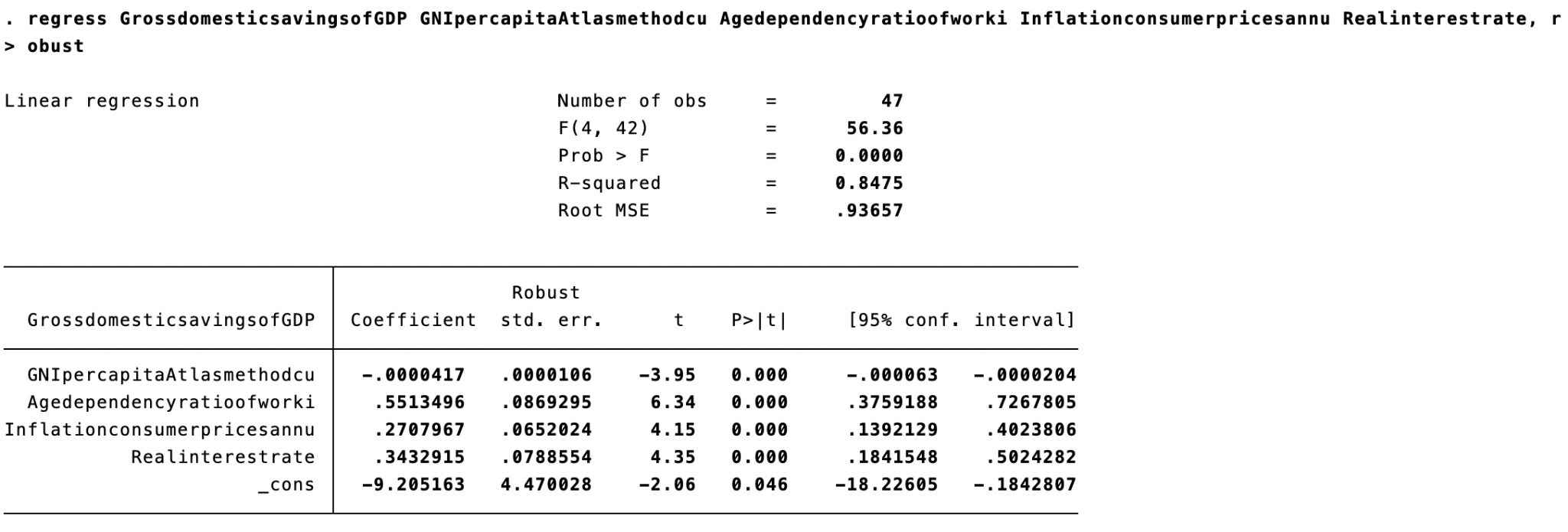
**Appendix 6.2: Breusch-Pagan test**

****

## **Appendix 6.3: The White’s test**

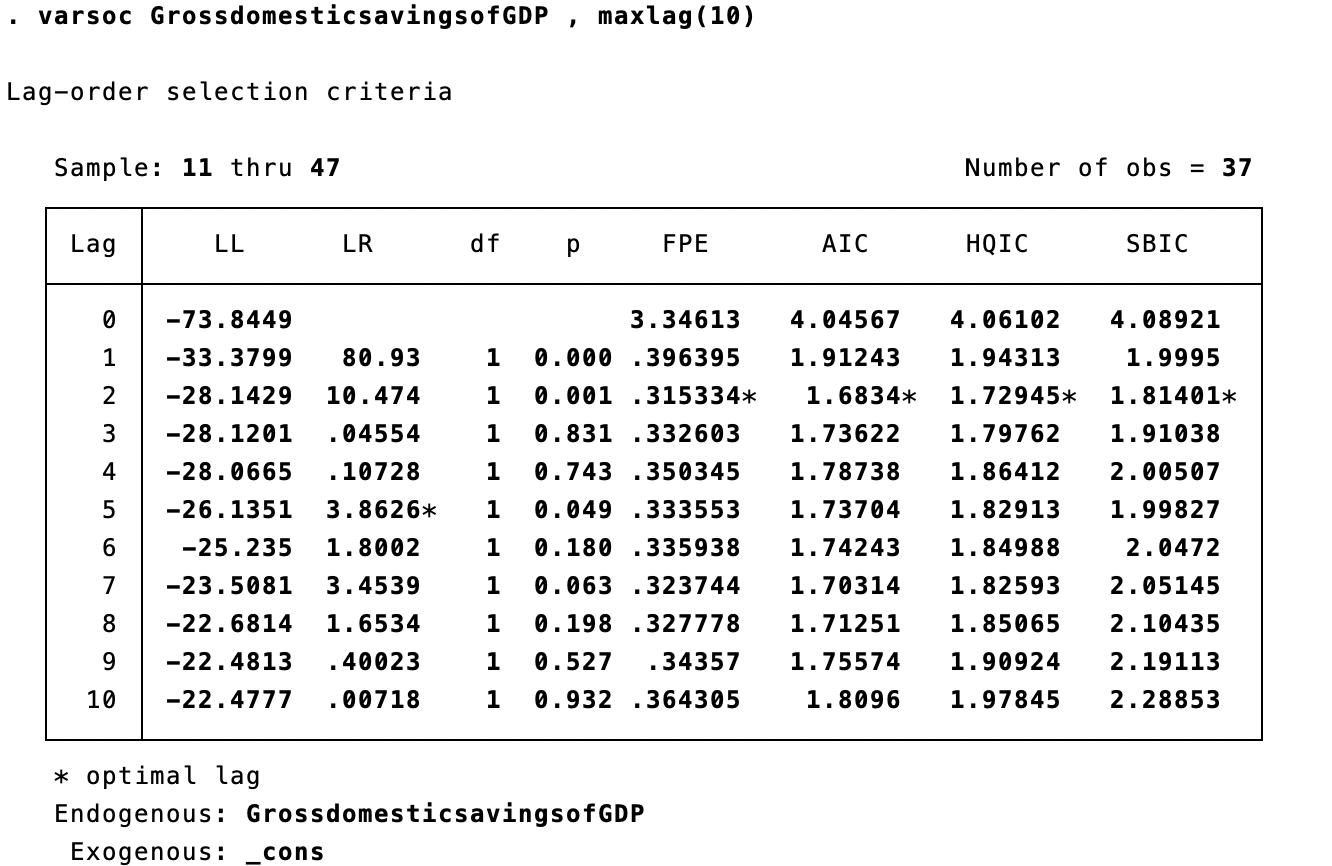
****

**Appendix 6.4: Robust test**

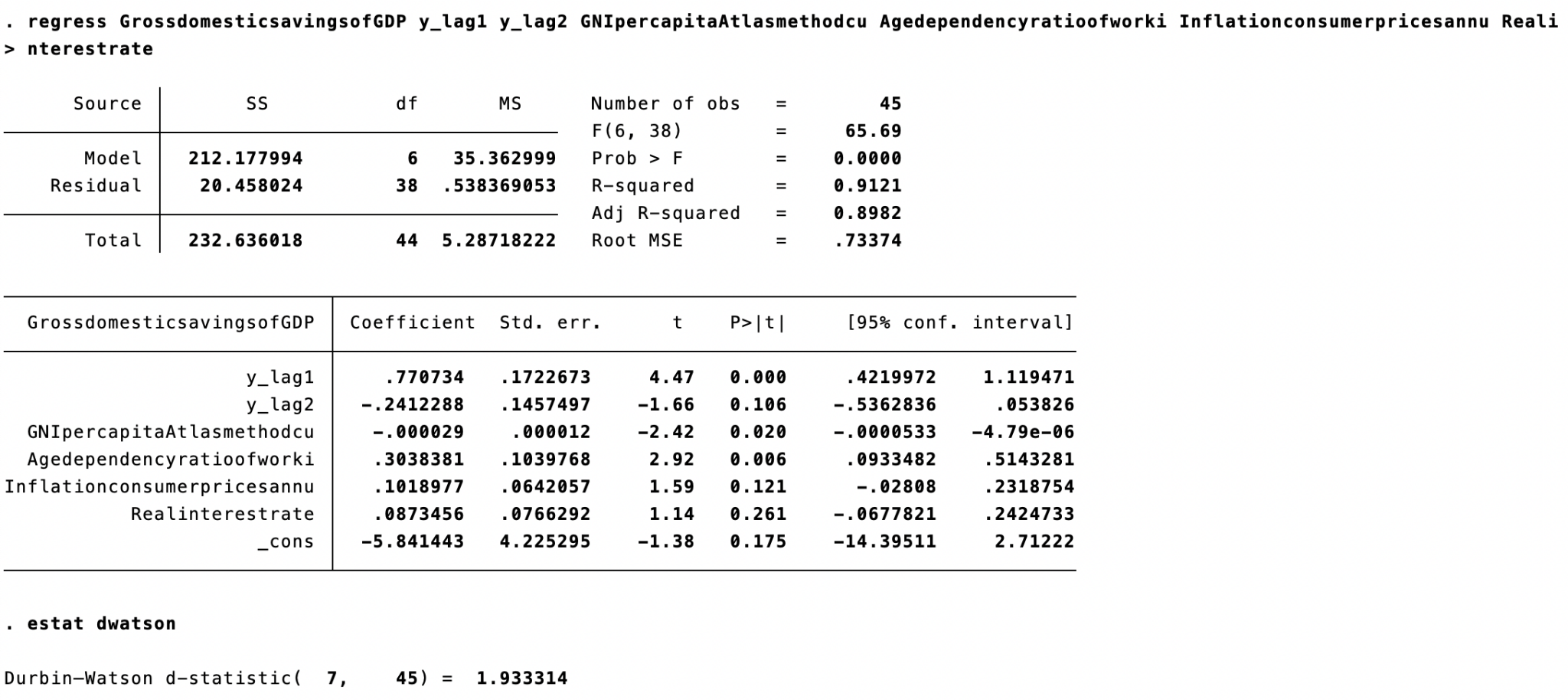
****

## **Appendix 7 : Durbin-Watson test**

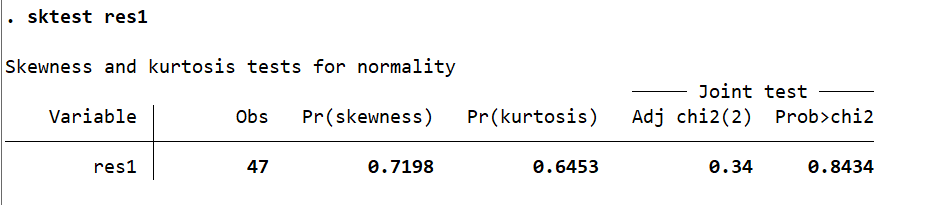
* **Appendix 7.1**



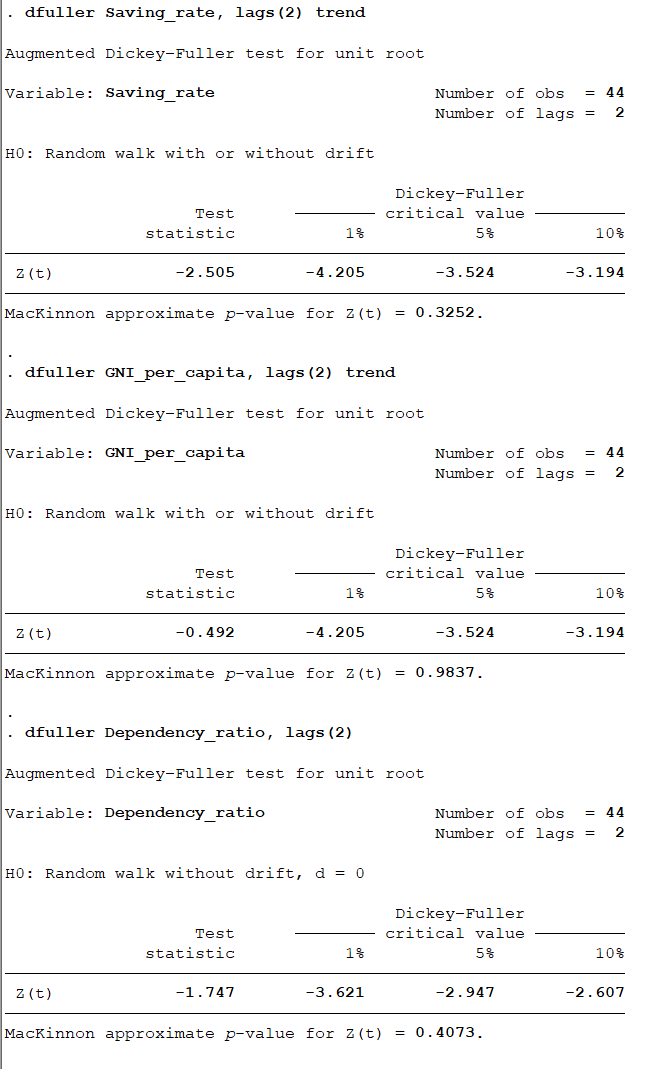
* **Appendix 7.2**

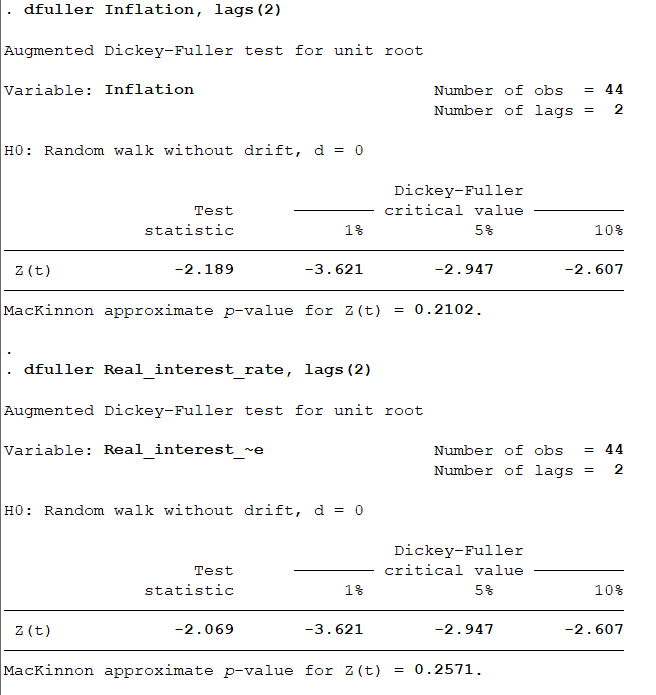
****

## **Appendix 8 : Normality**

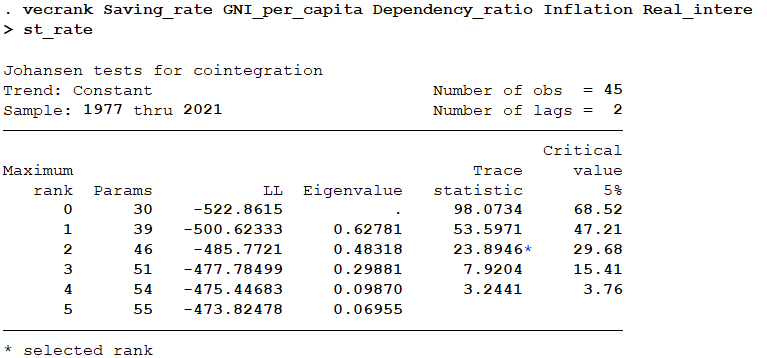


## **Appendix 9: Augmented Dickey-Fuller test**





**Appendix 10: The Johansen test**



**Appendix 11: VECM model**

