# Department of Economics

UNIVERSITY OF BATH

**STUDENT & UNIT DETAILS –** TO BE COMPLETED BY THE STUDENT

| | |
|---|---|
| **Candidate Number** | 21856 |
| **Unit Name and Code** | ES20069 and ES20159 (Introduction to) econometrics |

**DECLARATION –** PLEASE READ CAREFULLY

When you enrolled as a student at the University of Bath, you agreed to abide by the University's rules and regulations and agreed that you would access and read your programme handbook. This handbook contains references to, and penalties for, unfair practices such as collusion, plagiarism, fabrication or falsification. The University's Quality Assurance Code of Practice, QA53 Examination and Assessment Offences, sets out the consequences of committing an offence and the penalties that might be applied.

**By submitting this assessment, you confirm that:**

1. You have not impersonated, or allowed yourself to be impersonated by, any person for the purposes of this assessment.
2. This assessment is your original work and no part of it has been copied from any other source except where due acknowledgement is made.
3. You have not previously submitted this work for any other unit/course.
4. You give permission for your assessment response to be reproduced, communicated, compared and archived for plagiarism detection, benchmarking or educational purposes.
5. You understand that plagiarism is the presentation of the work, idea or creation of another person as though it is your own. It is a form of cheating and is a very serious academic offence that may lead to disciplinary action.
6. No part of this assessment has been produced for, or communicated to, you by any other person.

| **MARK AND COMMENTS** – TO BE COMPLETED BY THE MARKER | |
|---|---|
| | |
| | MARK (%) |

**START YOUR ASSESSMENT HERE.**

**BEFORE YOUR ANSWERS, STATE EXPLICITLY WHICH PART / QUESTION / SUBQUESTION YOU ARE ATTEMPTING.**

**1a.**

This is an AR(1) model where the stationary condition ($| \phi | < 1$) is not met as $\phi = 1$. Therefore it is not covariance stationary.

**1b.**

This is an ARMA(1,1) model where $\phi = 0.7$ , $\theta = 1$ . This model is covariance stationary if $| z | > 1$
In this case $\theta + \phi * z = 0$ $therefore$ $z = -\frac{\theta}{\phi} = -\frac{10}{7}$ so this model is covariance stationary.

**1c.**

The first axiom of if something is covariance stationary is that E($Y_t$) is constant however in this case E($Y_t$) = $\frac{t}{2}$ and is therefore time dependant and not constant. Thus it is not covariance stationary.

**2.**

$Y_t$ is an AR(1) model and the stationary condition is met, therefore it is covariance stationary.
The expected value is calculated as follows.

$$\text{E}(Y_t) = \text{E}(\alpha) + \phi * E\left(Y_{(t-1)}\right) + \epsilon_t = \alpha + \phi * E\left(Y_{(t-1)}\right) = \alpha + \phi * (\alpha + \phi * E(Y_{(t-2)}))$$

We can see that this pattern will repeat infinitely and therefore,

$$\text{E}(Y_t) = \lim_{n \to +\infty} \left( \left[\sum_{i=0}^{n} \phi^i * \alpha\right] + \phi^n * E\left(Y_{(t-n)}\right) \right) = \lim_{n \to +\infty} \left( \left[\sum_{i=0}^{n} \phi^i * \alpha\right] \right) = \frac{\alpha}{1-\phi}$$

Thus statement B is the only true statement.

**3a.**

The conditional mean of $Y_t$ is the mean of $Y_t$ given its previous terms. The previous terms are irrelevant however as E($Y_t$ | $Y_{(t-1)}, Y_{(t-2)}$ ... ) = $E(\sigma_t) * E(z_t)$ and $E(z_t) = 0$ thus the conditional mean, E($Y_t$ | $Y_{(t-1)}, Y_{(t-2)}$ ... ) = 0.

**3b.**

Since E($Y_t$) = 0, the standardised residuals reduces down to,

$$\frac{Y_t}{\sqrt{Var(Y_t)}}$$

$$Var(Y_t) = Var(\sigma_t * z_t) = Var(z_t)Var(\sigma_t) + Var(z_t)\left(E(\sigma_t)\right)^2 + Var(\sigma_t)\left(E(z_t)\right)^2$$
$$= Var(\sigma_t) + E(\sigma_t)^2 = E(\sigma_t^2)$$

Thus the standardised residuals are only dependant on $Y_t$ and 2 of its previous terms so this is an AR(2) model.

**4a.**

To find the existence of a unit root we can use the Dicky Fuller test. When looking at an AR(1)

$$Y_t = \rho * Y_{(t-1)} + u_t$$

A unit root is present if $\rho = 1$
Using this test we can put $Z_t$ and $X_t$ in the form of an AR(1) and look at the ρ value.

$$Z_t = 2 * Y_t + w_t = 2 * \left(Y_{(t-1)} + \epsilon_t\right) + w_t = 2 * \left(Y_{(t-2)} + \epsilon_{(t-1)} + \epsilon_t\right) + w_t$$
$$Z_{(t-1)} = 2 * Y_{(t-1)} + w_{(t-1)} = 2 * \left(Y_{(t-2)} + \epsilon_{(t-1)}\right) + w_{(t-1)}$$

By substitution we get that

$$Z_t = Z_{(t-1)} + u_t$$
$$where\ u_t = 2 * \epsilon_t + w_t - w_{(t-1)}$$

Thus $\rho = 1$ so $Z_t$ has a unit root.

$$X_t = Y_t + v_t = Y_{(t-1)} + \epsilon_t + v_t$$
$$X_{(t-1)} = Y_{(t-1)} + v_{(t-1)}$$

By substitution we get that

$$X_t = X_{(t-1)} + u_t$$
$$where\ u_t = \epsilon_t + v_t - v_{(t-1)}$$

Thus $\rho = 1$ so $X_t$ has a unit root.

**4b.**

Cointegration is the process of looking at a combination of non-stationary variables to see if it yields something stationary, this must be linear combination.

For $Y_t$ and $Z_t$ we can see that $Z_t - 2 * Y_t = w_t = N(0,1)$ which is covariance stationary so the cointegrating factors are 1 and -2.

For $X_t$ and $Z_t$ we can see that $Z_t - 2 * X_t = w_t - 2 * v_t = N(0,1) - 2 * N(0,1)$ which is also covariance stationary with cointegrating factors of 1 and -2.

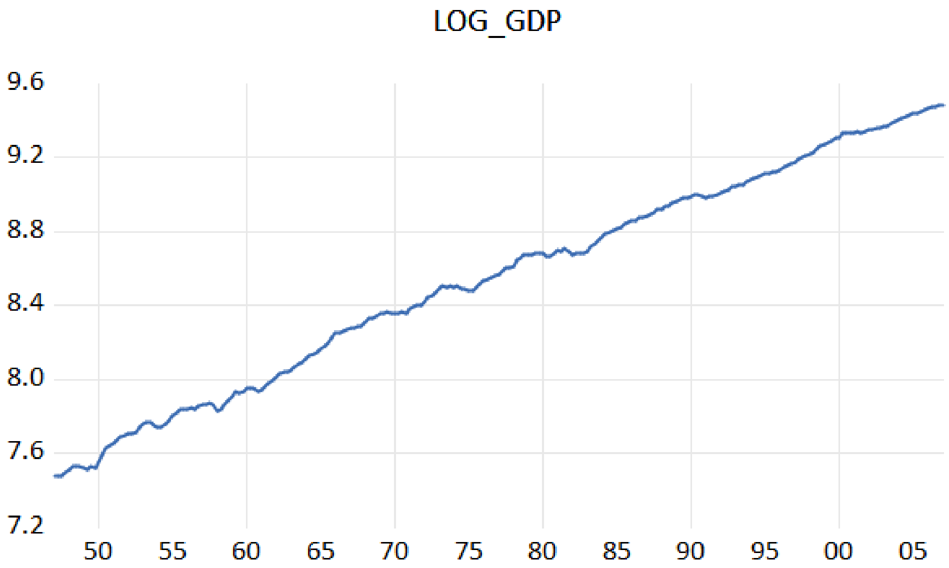
**4c.**

An error correction model is a dynamic model where the change of the variable is related to the distance between its value in the pervious period and its value in the long-run equilibrium. The error correction model in this case can be written as,

$$\Delta Y_t = \alpha_1\left(Z_{(t-1)} - \beta * Y_{(t-1)}\right) + \epsilon_t$$
$$\Delta Z_t = -\alpha_2\left(Z_{(t-1)} - \beta * Y_{(t-1)}\right) + w_t$$

Since $\Delta Y_t = Y_t - Y_{(t-1)} = \epsilon_t$ it implies that $\alpha_1 = 0$
Similarly $\Delta Z_t = Z_t - Z_{(t-1)} = 2\epsilon_t + w_t - w_{(t-1)}$ implies that $\alpha_2 = 1 - \frac{2\epsilon_t}{w_{(t-1)}}$

**5a.**




**5b.**

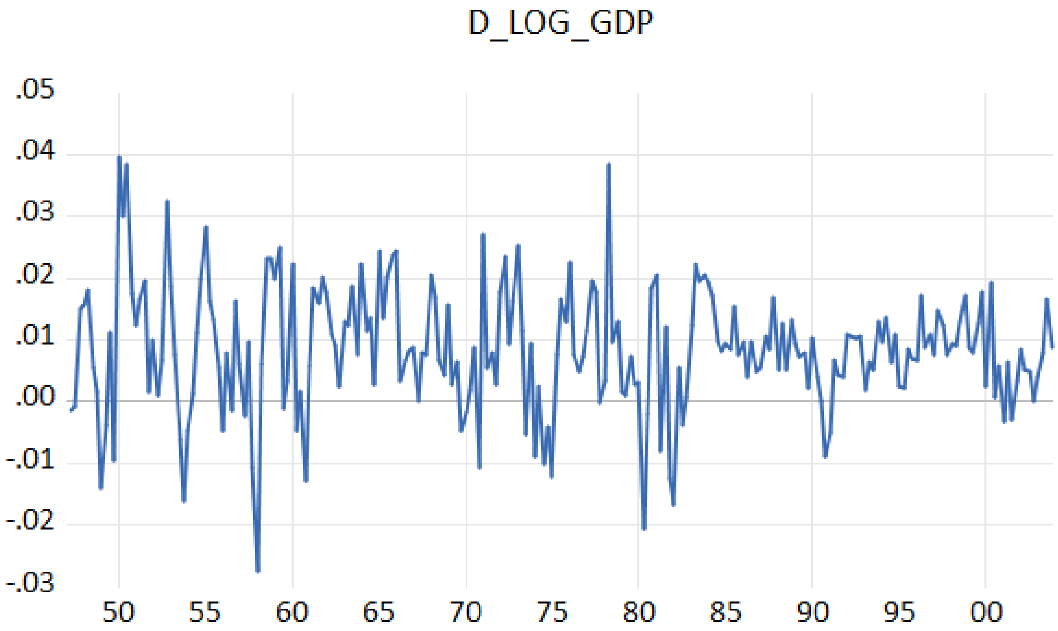
Date: 04/22/21 Time: 01:10
Sample: 1947Q1 2007Q1
Included observations: 241

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.987 | 0.987 | 237.90 | 0.000 |
| | | 2 | 0.974 | -0.018 | 470.59 | 0.000 |
| | | 3 | 0.961 | -0.014 | 698.02 | 0.000 |
| | | 4 | 0.948 | -0.006 | 920.25 | 0.000 |
| | | 5 | 0.935 | -0.004 | 1137.3 | 0.000 |
| | | 6 | 0.923 | 0.003 | 1349.4 | 0.000 |
| | | 7 | 0.910 | -0.009 | 1556.6 | 0.000 |
| | | 8 | 0.897 | -0.010 | 1758.7 | 0.000 |
| | | 9 | 0.884 | -0.020 | 1955.8 | 0.000 |
| | | 10 | 0.870 | -0.013 | 2147.8 | 0.000 |
| | | 11 | 0.857 | -0.006 | 2334.8 | 0.000 |
| | | 12 | 0.843 | -0.017 | 2516.6 | 0.000 |
| | | 13 | 0.830 | 0.010 | 2693.6 | 0.000 |
| | | 14 | 0.817 | 0.005 | 2866.0 | 0.000 |
| | | 15 | 0.805 | 0.010 | 3033.8 | 0.000 |
| | | 16 | 0.793 | 0.003 | 3197.4 | 0.000 |
| | | 17 | 0.781 | -0.005 | 3356.8 | 0.000 |
| | | 18 | 0.769 | -0.004 | 3511.9 | 0.000 |
| | | 19 | 0.757 | -0.005 | 3662.9 | 0.000 |
| | | 20 | 0.745 | -0.011 | 3809.8 | 0.000 |
| | | 21 | 0.732 | -0.008 | 3952.6 | 0.000 |
| | | 22 | 0.720 | -0.009 | 4091.3 | 0.000 |
| | | 23 | 0.708 | -0.009 | 4226.0 | 0.000 |
| | | 24 | 0.696 | 0.002 | 4356.8 | 0.000 |
| | | 25 | 0.684 | -0.001 | 4483.7 | 0.000 |
| | | 26 | 0.672 | -0.012 | 4606.8 | 0.000 |
| | | 27 | 0.660 | -0.015 | 4726.1 | 0.000 |
| | | 28 | 0.648 | -0.022 | 4841.4 | 0.000 |
| | | 29 | 0.635 | -0.008 | 4952.8 | 0.000 |
| | | 30 | 0.622 | -0.013 | 5060.3 | 0.000 |
| | | 31 | 0.610 | 0.001 | 5164.0 | 0.000 |
| | | 32 | 0.598 | 0.002 | 5264.0 | 0.000 |
| | | 33 | 0.586 | 0.004 | 5360.6 | 0.000 |
| | | 34 | 0.574 | -0.002 | 5453.7 | 0.000 |
| | | 35 | 0.562 | -0.001 | 5543.6 | 0.000 |
| | | 36 | 0.551 | -0.005 | 5630.2 | 0.000 |

It is an AR(1) model where $Y_t = 0.987 * Y_{(t-1)} + u_t$. This is because of the high level of autocorrelation and the fact that each term is approximately $0.987^{\mathrm{k}}$ where k is the index of the element.

**5c.**

We can follow the 7 step Box-Jenkins process to estimate and forecast the "best: ARMA model(s)
STEP 1 – Apply data transformation if necessary. Test for stationarity/nonstationary. If the series is nonstationary, transform the data to achieve stationarity.
As shown by part a, log(gdp) is nonstationary so we may apply the transformation of the 1st difference.




This initially appears to be stationary so we can test it using ADF and KPSS

Null Hypothesis: D_LOG_GDP has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

| | | t-Statistic | Prob.* |
|---|---|---|---|
| Augmented Dickey-Fuller test statistic | | -10.59543 | 0.0000 |
| Test critical values: | 1% level | -3.459231 | |
| | 5% level | -2.874143 | |
| | 10% level | -2.573563 | |

ADF shows us that it is stationary as the absolute value of the test-statistic is greater than the absolute value of the critical value at a 5% level, therefore we reject the null hypothesis of it being non-stationary.

Null Hypothesis: D_LOG_GDP is stationary
Exogenous: Constant
Bandwidth: 2 (Newey-West automatic) using Bartlett kernel

| | | LM-Stat. |
|---|---|---|
| Kwiatkowski-Phillips-Schmidt-Shin test statistic | | 0.107492 |
| Asymptotic critical values*: | 1% level | 0.739000 |
| | 5% level | 0.463000 |
| | 10% level | 0.347000 |

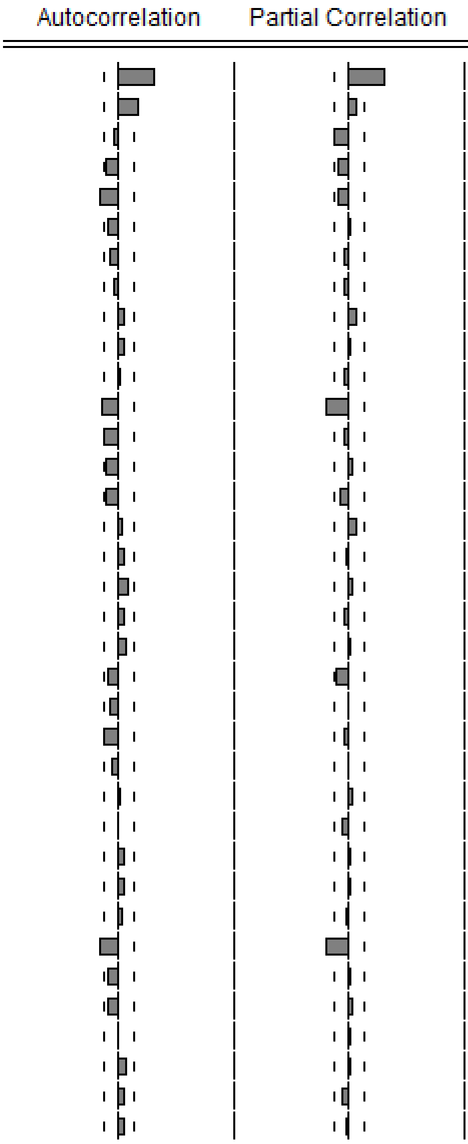
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

| | |
|---|---|
| Residual variance (no correction) | 0.000101 |
| HAC corrected variance (Bartlett kernel) | 0.000159 |

KPSS also shows us that it is stationary as the absolute value of the LM-Statistic is less than the absolute value of the critical value at a 5% level, therefore we accept the null hypothesis that it is stationary.

STEP 2 – Determine possible lag orders p and q in the ARMA(p,q) using (PACF/ACF)

Date: 04/22/21 Time: 03:24
Sample: 1947Q1 2003Q4
Included observations: 227




| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | 0.334 | 0.334 | 25.729 | 0.000 |
| 2 | 0.181 | 0.078 | 33.276 | 0.000 |
| 3 | -0.030 | -0.126 | 33.479 | 0.000 |
| 4 | -0.114 | -0.096 | 36.496 | 0.000 |
| 5 | -0.169 | -0.094 | 43.197 | 0.000 |
| 6 | -0.085 | 0.024 | 44.880 | 0.000 |
| 7 | -0.070 | -0.029 | 46.022 | 0.000 |
| 8 | -0.040 | -0.035 | 46.394 | 0.000 |
| 9 | 0.056 | 0.072 | 47.149 | 0.000 |
| 10 | 0.059 | 0.011 | 47.990 | 0.000 |
| 11 | 0.019 | -0.041 | 48.081 | 0.000 |
| 12 | -0.146 | -0.189 | 53.231 | 0.000 |
| 13 | -0.135 | -0.043 | 57.685 | 0.000 |
| 14 | -0.100 | 0.027 | 60.129 | 0.000 |
| 15 | -0.101 | -0.072 | 62.648 | 0.000 |
| 16 | 0.041 | 0.075 | 63.053 | 0.000 |
| 17 | 0.052 | -0.009 | 63.725 | 0.000 |
| 18 | 0.096 | 0.035 | 66.032 | 0.000 |
| 19 | 0.051 | -0.032 | 66.684 | 0.000 |
| 20 | 0.068 | 0.009 | 67.838 | 0.000 |
| 21 | -0.094 | -0.112 | 70.062 | 0.000 |
| 22 | -0.070 | 0.004 | 71.310 | 0.000 |
| 23 | -0.118 | -0.045 | 74.845 | 0.000 |
| 24 | -0.051 | -0.003 | 75.519 | 0.000 |
| 25 | 0.015 | 0.031 | 75.577 | 0.000 |
| 26 | 0.004 | -0.064 | 75.582 | 0.000 |
| 27 | 0.060 | 0.015 | 76.509 | 0.000 |
| 28 | 0.053 | 0.019 | 77.235 | 0.000 |
| 29 | 0.025 | -0.021 | 77.402 | 0.000 |
| 30 | -0.158 | -0.197 | 83.959 | 0.000 |
| 31 | -0.095 | 0.021 | 86.330 | 0.000 |
| 32 | -0.086 | 0.036 | 88.288 | 0.000 |
| 33 | 0.006 | 0.025 | 88.299 | 0.000 |
| 34 | 0.067 | 0.010 | 89.512 | 0.000 |
| 35 | 0.051 | -0.061 | 90.224 | 0.000 |
| 36 | 0.046 | -0.017 | 90.810 | 0.000 |

From this we can drive values for p using Partial Correlation and q using Autocorrelation. Looking at the patterns in the ACF and PACF a model of ARMA(p,q) seems most suitable as all other models do not fit the shape of the functions.

STEP 3 – Estimate the tentative models identified in step 2

Using this model P = (0,1) and Q = (0,1,2) leading to 5 possible combinations shown in step 4

STEP 4 – Compare and estimate models using an information criterion.

My two information criterions being used is AIC and SBC, below is the table showing the possible models with the lower score being better.

| ARMA(p,q) | AIC | SBC |
|---|---|---|
| ARMA(0,1) | -6.423458 | -6.378195 |
| ARMA(0,2) | -6.376201 | -6.330937 |
| ARMA(1,0) | -6.450735 | **-6.405471** |
| ARMA(1,1) | -6.445264 | -6.384913 |
| ARMA(1,2) | **-6.458009** | -6.397658 |

Thus we have identified two possible models, AR(1) and ARMA(1,2)

STEP 5 – Test for autocorrelation in the error terms

As both models have p values of one we can use Durbin-Watson test. This test tell us that there is no autocorrelation when the statistic is close to 2.

For the AR(1) model we can use the Durbin-Watson statistic to test for autocorrelation. The DW statistics is 2.047677 which implies that there is no autocorrelation.

For the ARMA(1,2) we can also use the Durbin-Watson statistic to test for autocorrelation. The DW statistics is 1.975754 which implies that there is no autocorrelation.

STEP 6 – Test for heteroscedasticity (Testing for ARCH effects)

For AR(1):

Heteroskedasticity Test: ARCH

| | | | |
|---|---|---|---|
| F-statistic | 2.529485 | Prob. F(1,224) | 0.1131 |
| Obs*R-squared | 2.523573 | Prob. Chi-Square(1) | 0.1122 |

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/22/21 Time: 03:35
Sample (adjusted): 1947Q3 2003Q4
Included observations: 226 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 8.05E-05 | 1.27E-05 | 6.346576 | 0.0000 |
| RESID^2(-1) | 0.105731 | 0.066479 | 1.590435 | 0.1131 |

| | | | |
|---|---|---|---|
| R-squared | 0.011166 | Mean dependent var | 9.00E-05 |
| Adjusted R-squared | 0.006752 | S.D. dependent var | 0.000168 |
| S.E. of regression | 0.000168 | Akaike info criterion | -14.53840 |
| Sum squared resid | 6.31E-06 | Schwarz criterion | -14.50813 |
| Log likelihood | 1644.840 | Hannan-Quinn criter. | -14.52619 |
| F-statistic | 2.529485 | Durbin-Watson stat | 2.029710 |
| Prob(F-statistic) | 0.113147 | | |

We do not reject the null hypothesis as Prob.F and Prob.Chi-Square are greater than 5% and thus there is no conditional heteroscedasticity.

For ARMA(1,2)

Heteroskedasticity Test: ARCH

| | | | |
|---|---|---|---|
| F-statistic | 3.130271 | Prob. F(1,224) | 0.0782 |
| Obs*R-squared | 3.114694 | Prob. Chi-Square(1) | 0.0776 |

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/22/21 Time: 03:36
Sample (adjusted): 1947Q3 2003Q4
Included observations: 226 after adjustments

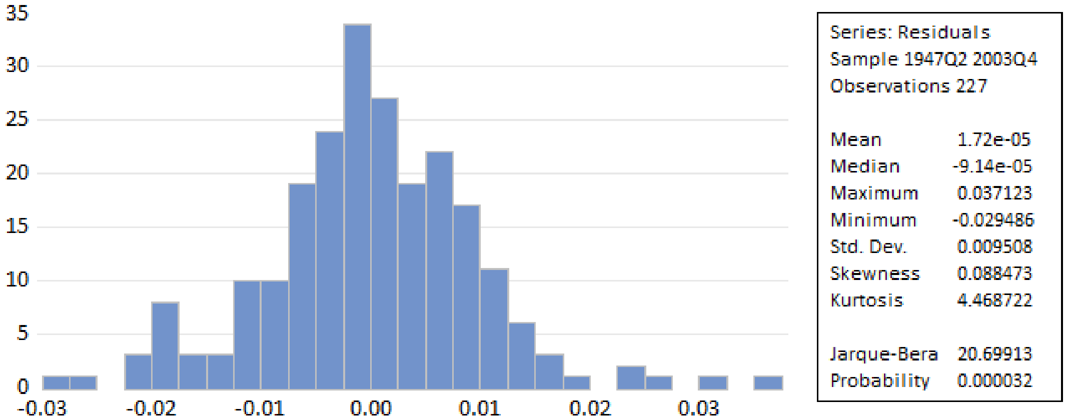
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 7.81E-05 | 1.25E-05 | 6.258828 | 0.0000 |
| RESID^2(-1) | 0.117462 | 0.066391 | 1.769257 | 0.0782 |

| | | | |
|---|---|---|---|
| R-squared | 0.013782 | Mean dependent var | 8.86E-05 |
| Adjusted R-squared | 0.009379 | S.D. dependent var | 0.000166 |
| S.E. of regression | 0.000165 | Akaike info criterion | -14.56838 |
| Sum squared resid | 6.12E-06 | Schwarz criterion | -14.53811 |
| Log likelihood | 1648.227 | Hannan-Quinn criter. | -14.55617 |
| F-statistic | 3.130271 | Durbin-Watson stat | 2.035892 |
| Prob(F-statistic) | 0.078212 | | |

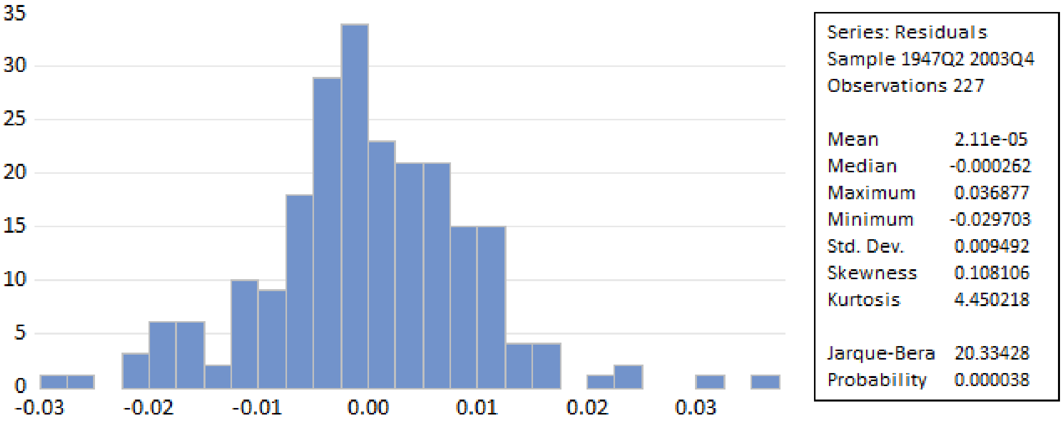
We do not reject the null hypothesis as Prob.F and Prob.Chi-Square are greater than 5% and thus there is no conditional heteroscedasticity.

STEP 7 – Test for normality (Jarque – Bera test)

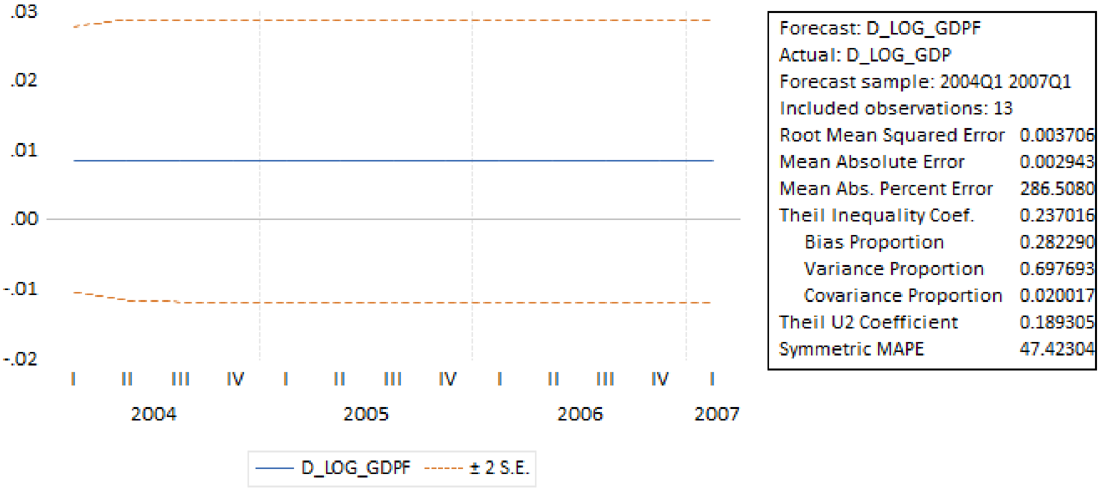
To test normality we can use the Jarque-Bera test statistic and check is the P-Value is less than 5%
For AR(1)



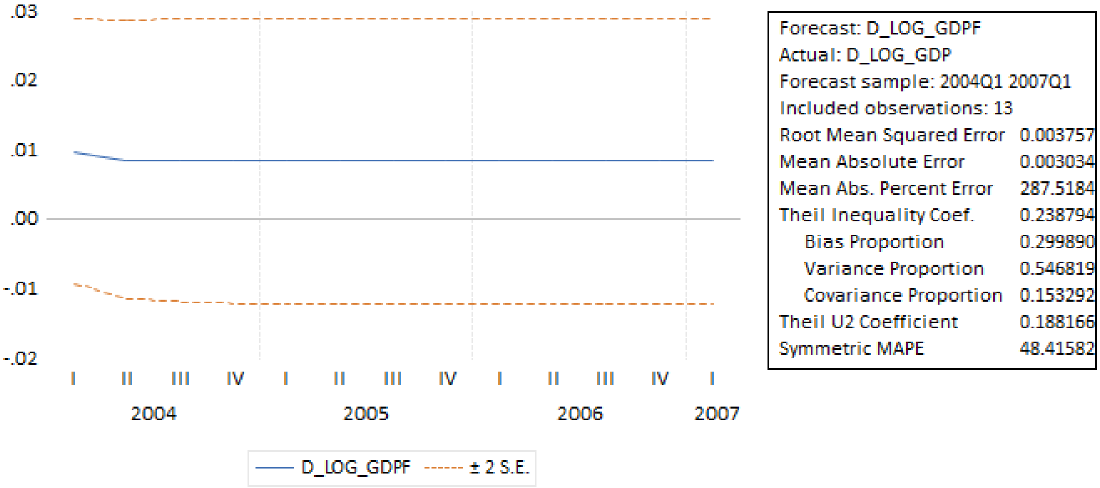

As shown by the distribution and P-Value below 5%, the errors are normally distributed.
For ARMA(1,2)




This model also has its errors normally distributed as the P-Value is below 5%.
FORCASTING
AR(1)




ARMA(1,2)




Both the AR(1) and ARMA(1,2) have very low MSE and MAE therefore we can conclude that they fit the data to a sufficient extent and is therefore a good model. However AR(1) appears to be slightly more accurate and therefore is the “best” ARMA model.

Dependent Variable: D_LOG_GDP
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/22/21 Time: 04:24
Sample: 1947Q2 2003Q4
Included observations: 227
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.008421 | 0.000962 | 8.754657 | 0.0000 |
| AR(1) | 0.334434 | 0.055834 | 5.989739 | 0.0000 |
| SIGMASQ | 9.00E-05 | 6.45E-06 | 13.94813 | 0.0000 |

| | | | |
|---|---|---|---|
| R-squared | 0.112345 | Mean dependent var | 0.008442 |
| Adjusted R-squared | 0.104420 | S.D. dependent var | 0.010092 |
| S.E. of regression | 0.009551 | Akaike info criterion | -6.450735 |
| Sum squared resid | 0.020433 | Schwarz criterion | -6.405471 |
| Log likelihood | 735.1584 | Hannan-Quinn criter. | -6.432471 |
| F-statistic | 14.17517 | Durbin-Watson stat | 2.047677 |
| Prob(F-statistic) | 0.000002 | | |

| | |
|---|---|
| Inverted AR Roots | .33 |

This is a summary of the model and therefore we can calculate the growth from the first difference. Growth = $e^{0.008421+0.334434*(firstdifferance)}$ = 0.40879. Therefore, the growth per quarter is 0.40879

**6a.**

To access stationarity for growth in log consumption and log income we will use an ADF test.

Null Hypothesis: D_CON has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on SIC, maxlag=12)

| | | t-Statistic | Prob.* |
|---|---|---|---|
| Augmented Dickey-Fuller test statistic | | -2.837276 | 0.0563 |
| Test critical values: | 1% level | -3.488585 | |
| | 5% level | -2.886959 | |
| | 10% level | -2.580402 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(D_CON)
Method: Least Squares
Date: 04/22/21 Time: 05:34
Sample (adjusted): 1987Q1 2015Q2
Included observations: 114 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| D_CON(-1) | -0.281437 | 0.099193 | -2.837276 | 0.0054 |
| D(D_CON(-1)) | -0.517427 | 0.103673 | -4.990939 | 0.0000 |
| D(D_CON(-2)) | -0.346554 | 0.087543 | -3.958656 | 0.0001 |
| C | 0.001890 | 0.000800 | 2.363015 | 0.0199 |

| | | | |
|---|---|---|---|
| R-squared | 0.430694 | Mean dependent var | 2.10E-05 |
| Adjusted R-squared | 0.415168 | S.D. dependent var | 0.005668 |
| S.E. of regression | 0.004335 | Akaike info criterion | -8.009823 |
| Sum squared resid | 0.002067 | Schwarz criterion | -7.913816 |
| Log likelihood | 460.5599 | Hannan-Quinn criter. | -7.970859 |
| F-statistic | 27.73925 | Durbin-Watson stat | 1.885390 |
| Prob(F-statistic) | 0.000000 | | |

Null Hypothesis: D_CON has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on SIC, maxlag=12)

| | | t-Statistic | Prob.* |
|---|---|---|---|
| Augmented Dickey-Fuller test statistic | | -2.837276 | 0.0563 |
| Test critical values: | 1% level | -3.488585 | |
| | 5% level | -2.886959 | |
| | 10% level | -2.580402 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(D_CON)
Method: Least Squares
Date: 04/22/21 Time: 05:28
Sample (adjusted): 1987Q1 2015Q2
Included observations: 114 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| D_CON(-1) | -0.281437 | 0.099193 | -2.837276 | 0.0054 |
| D(D_CON(-1)) | -0.517427 | 0.103673 | -4.990939 | 0.0000 |
| D(D_CON(-2)) | -0.346554 | 0.087543 | -3.958656 | 0.0001 |
| C | 0.001890 | 0.000800 | 2.363015 | 0.0199 |

| | | | |
|---|---|---|---|
| R-squared | 0.430694 | Mean dependent var | 2.10E-05 |
| Adjusted R-squared | 0.415168 | S.D. dependent var | 0.005668 |
| S.E. of regression | 0.004335 | Akaike info criterion | -8.009823 |
| Sum squared resid | 0.002067 | Schwarz criterion | -7.913816 |
| Log likelihood | 460.5599 | Hannan-Quinn criter. | -7.970859 |
| F-statistic | 27.73925 | Durbin-Watson stat | 1.885390 |
| Prob(F-statistic) | 0.000000 | | |

We can see that the null hypothesis is rejected both times so both are stationary.
We then build the model and determine the optimal number of lags which turns out to be 4 as seen in the chart below.

VAR Lag Order Selection Criteria
Endogenous variables: D_CON D_INC
Exogenous variables: C
Date: 04/22/21 Time: 05:26
Sample: 1986Q1 2015Q2
Included observations: 109

| Lag | LogL | LR | FPE | AIC | SC | HQ |
|---|---|---|---|---|---|---|
| 0 | 787.4265 | NA | 1.89e-09 | -14.41150 | -14.36211 | -14.39147 |
| 1 | 813.0058 | 49.75054 | 1.27e-09 | -14.80745 | -14.65930* | -14.74737* |
| 2 | 818.6724 | 10.81333 | 1.23e-09 | -14.83803 | -14.59111 | -14.73789 |
| 3 | 825.0099 | 11.86102* | 1.18e-09 | -14.88092 | -14.53524 | -14.74073 |
| 4 | 829.3434 | 7.951369 | 1.17e-09* | -14.88704* | -14.44259 | -14.70680 |
| 5 | 829.8570 | 0.923596 | 1.25e-09 | -14.82307 | -14.27986 | -14.60277 |
| 6 | 830.8932 | 1.825149 | 1.32e-09 | -14.76868 | -14.12671 | -14.50834 |
| 7 | 833.7180 | 4.872188 | 1.36e-09 | -14.74712 | -14.00638 | -14.44672 |
| 8 | 836.2297 | 4.239870 | 1.39e-09 | -14.71981 | -13.88031 | -14.37936 |

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Now that we have re-estimated the VAR model we have to perform the necessary diagnostic tests. The three tables below outline tests for autocorrelation, heteroskedasticity and normality and show us that there us no autocorrelation, no heteroskedasticity and no normality.

VAR Residual Serial Correlation LM Tests
Date: 04/22/21 Time: 05:39
Sample: 1986Q1 2015Q2
Included observations: 113

Null hypothesis: No serial correlation at lag h

| Lag | LRE* stat | df | Prob. | Rao F-stat | df | Prob. |
|---|---|---|---|---|---|---|
| 1 | 4.959442 | 4 | 0.2915 | 1.248947 | (4, 202.0) | 0.2915 |
| 2 | 6.506540 | 4 | 0.1644 | 1.644841 | (4, 202.0) | 0.1644 |
| 3 | 6.928773 | 4 | 0.1397 | 1.753414 | (4, 202.0) | 0.1397 |
| 4 | 1.294464 | 4 | 0.8623 | 0.323051 | (4, 202.0) | 0.8623 |

Null hypothesis: No serial correlation at lags 1 to h

| Lag | LRE* stat | df | Prob. | Rao F-stat | df | Prob. |
|---|---|---|---|---|---|---|
| 1 | 4.959442 | 4 | 0.2915 | 1.248947 | (4, 202.0) | 0.2915 |
| 2 | 9.396821 | 8 | 0.3099 | 1.184544 | (8, 198.0) | 0.3100 |
| 3 | 11.86761 | 12 | 0.4564 | 0.993447 | (12, 194.0) | 0.4566 |
| 4 | 15.17276 | 16 | 0.5120 | 0.950744 | (16, 190.0) | 0.5126 |

*Edgeworth expansion corrected likelihood ratio statistic.

VAR Residual Heteroskedasticity Tests (Levels and Squares)
Date: 04/22/21 Time: 05:40
Sample: 1986Q1 2015Q2
Included observations: 113

Joint test:

| Chi-sq | df | Prob. |
|---|---|---|
| 77.27782 | 48 | 0.0047 |

Individual components:

| Dependent | R-squared | F(16,96) | Prob. | Chi-sq(16) | Prob. |
|---|---|---|---|---|---|
| res1*res1 | 0.169069 | 1.220814 | 0.2667 | 19.10477 | 0.2633 |
| res2*res2 | 0.267193 | 2.187689 | 0.0102 | 30.19276 | 0.0170 |
| res2*res1 | 0.168375 | 1.214790 | 0.2712 | 19.02637 | 0.2673 |

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: Residuals are multivariate normal
Date: 04/22/21 Time: 05:41
Sample: 1986Q1 2015Q2
Included observations: 113

| Component | Skewness | Chi-sq | df | Prob.* |
|---|---|---|---|---|
| 1 | -0.046765 | 0.041187 | 1 | 0.8392 |
| 2 | -0.733088 | 10.12137 | 1 | 0.0015 |
| Joint | | 10.16256 | 2 | 0.0062 |

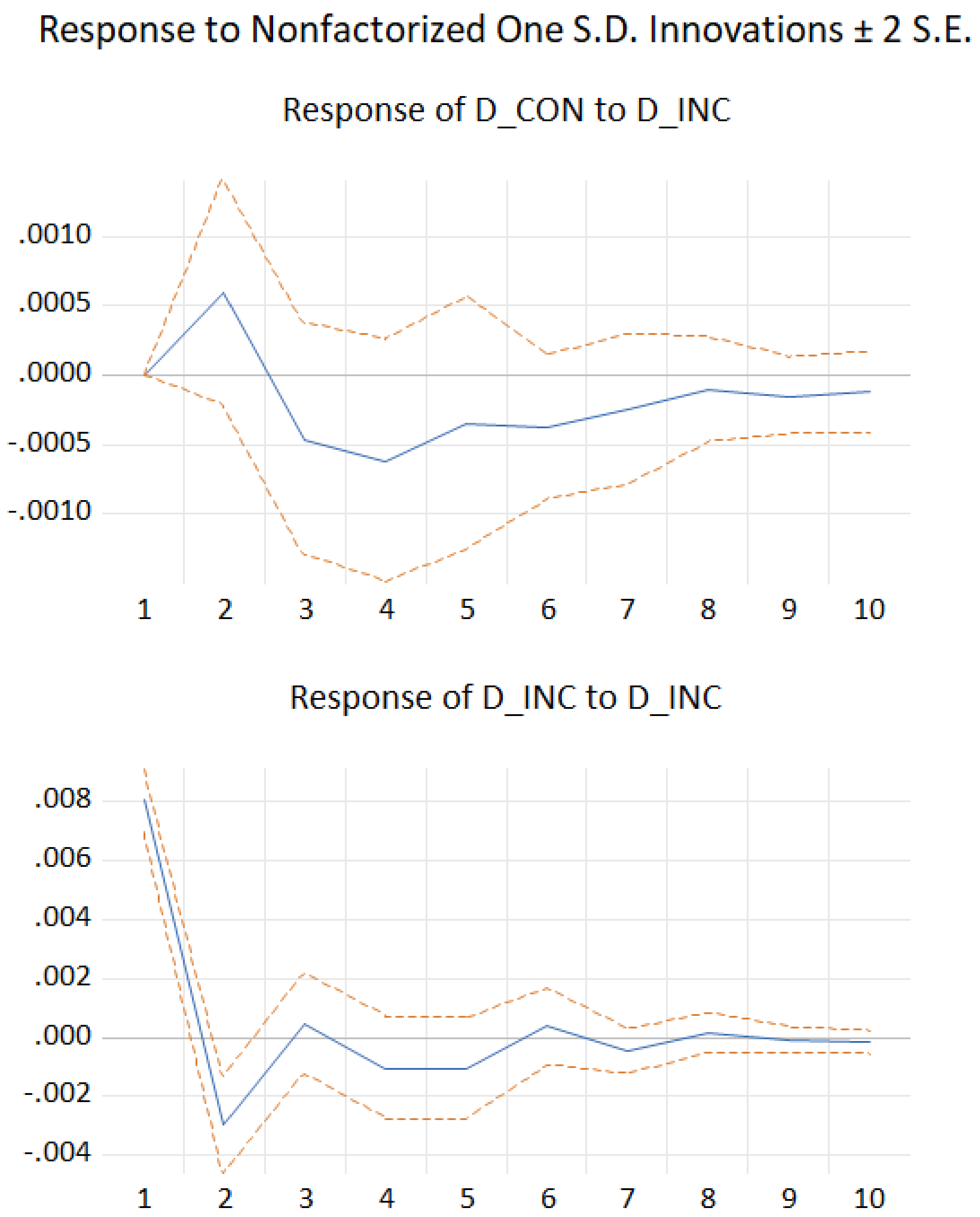
| Component | Kurtosis | Chi-sq | df | Prob. |
|---|---|---|---|---|
| 1 | 3.953530 | 4.280904 | 1 | 0.0385 |
| 2 | 7.515836 | 96.01597 | 1 | 0.0000 |
| Joint | | 100.2969 | 2 | 0.0000 |

| Component | Jarque-Bera | df | Prob. |
|---|---|---|---|
| 1 | 4.322091 | 2 | 0.1152 |
| 2 | 106.1373 | 2 | 0.0000 |
| Joint | 110.4594 | 4 | 0.0000 |

*Approximate p-values do not account for coefficient estimation

**6b.**

These following graphs show changes due to the log difference in income having a positive initial shock. With regards to consumption, there is an initial increase in period 2, then a larger drop in period 3 and 4 in which consumption never recovers to its original state. Income has an initial decline in period 2 and then fluctuates around and tends towards its original state.



**6c.**

Performing a Granger Causality test on the growth of log income and log consumption yielded this table

Pairwise Granger Causality Tests
Date: 04/22/21 Time: 06:21
Sample: 1986Q1 2015Q2
Lags: 4

| Null Hypothesis: | Obs | F-Statistic | Prob. |
|---|---|---|---|
| D_INC does not Granger Cause D_CON | 113 | 1.94465 | 0.1085 |
| D_CON does not Granger Cause D_INC | | 5.79399 | 0.0003 |

We are only interested if D_INC granger causes growth on D_CON however since the P-Value is greater than 5% we accept the null hypothesis that it does not cause growth.