1. Open the file WAGESASSESSMENT
2. Regress WAGE on IQ and report your results (2 marks)

WAGEi = β0 + β1IQi+ ui

H0: β1 = 0 (IQ does not affect wage), H1: β1 ≠ 0 (IQ effects wage)

R-squared = 0.095, meaning 9.5% of the variation in wage is explained by the variation in IQ, this suggests the explanatory power of the model is weak.

WAGEi = 117.597 + 8.298IQi+ ui

Sig = 0.000, < 0.05; IQ is statistically significant for a significance level of 5%, thus rejecting H0. For every 1 point increase in IQ, weekly wage increases by $8.30. β0 is 117.597, meaning the minimum weekly wage is $117.60.

1. Does the model in a) exhibit heteroscedasticity? Explain your answer (4 marks)

Heteroscedasticity is when the size of the error term differs across values of an independent variable. To determine whether a model exhibits heteroscedasticity, we run a breusch–pagan test. Within the test, we regress wage on IQ saving the unstandardized residuals. Secondly, squaring the residuals interprets the variance of disturbance, or error terms.

e2i= δ0 + δ1IQi + ui.

H0: δ1 = 0 (homoscedasticity)

H1: δ1 ≠ 0. (Heteroscedasticity)

Getting an f statistic value of 11.052 and a p value of 0.001, means we reject H0 for significance level of 5% (sig<0.05), hence the model exhibits heteroscedasticity.

1. Are the variables WAGE and IQ normally distributed? Justify your answer (4 marks)

When a Variable is normally distributed, all points lie on the 45-degree line on the Normal Q-Q plot. Wage is not normally distributed and is right-skewed. This can also be determined by the Normal Q-Q plot, a pair amount of points do not lie on the line, and as they start and end below it, wage is rightly skewed.

IQ is approximately normally distributed. On the Normal Q-Q plot, many points lie on the line or very close to it and the histogram matches up.

1. Are the residuals in your model normally distributed? Justify your answer (3 marks)

Residuals in this model are not normally distributed as shown by the histogram. The normal Q-Q plot supports this as many points do not lie on the line and they begin and end below it, indicating residuals are right-skewed.

1. Is the model in a) misspecified? Explain your answer (3 marks)

As the model is linear, the model will be misspecified if there is no linearity. To determine linearity, there must be symmetry in a scatter graph between the residuals and predicted values.

There is no symmetry through 0, therefore there is no linearity, hence the model is misspecified.

1. Open the file WAGESASSESSMENT
2. Regress LOGWAGE on LOGIQ and report your results (2 marks)

LogWagei = β0 + β1LogIQi+ ui

H0: β1 = 0 (LogIQ does not affect LogWage)

H1: β1 ≠ 0 (LogIQ effects LogWage)

R2 = 0.097; 9.7% of the variation in LogWage is explained by the variation in LogIQ, suggesting explanatory power of the model is weak.

LogWagei = 2.943 + 0.833LogIQi+ ui

LogIQ is statistically significant for 5% significance level (sig< 0.05) so Reject H0.

Every 1% increase in IQ points, increases weekly wage by 0.833%.

1. Does the model in a) exhibit heteroscedasticity? Explain your answer (4 marks)

As explained in 1b, we end up with an expression, e2 = δ0 + δ1LogIQ + ui.

H0: δ1 = 0 (Homoscedasticity)

H1: δ1 ≠ 0 (Heteroscedasticity).

A p-value larger than 0.05 means we cannot reject H0 of homoscedasticity for significance level of 5%. Meaning the model does not exhibit heteroscedasticity.

1. Are the variables LOGWAGE and LOGIQ normally distributed? Justify your answer (4 marks)

LogWage is approximately normally distributed as shown by the histogram. The Normal Q-Q plot supports this as many points lie on the line.

LogIQ isn’t normally distributed and the data is left-skewed, indicated by both the histogram and the Normal Q-Q plot. Many of the points do not lie on the line, showing LogIQ isn’t normally distributed. The points start as well as end above the line, meaning LogIQ is left-skewed.

1. Are the residuals in your model normally distributed? Justify your answer (3 marks)

The residuals are approximately normally distributed, represented by the histogram and the normal Q-Q plot. Many of the points lie on the line, showing the residuals are approximately normally distributed.

1. Is the model in a) mispecified? Explain your answer (3 marks)

As this model is linear, the model will be misspecified if no linearity is present. This will be determined by symmetry in a scatter graph between the residuals and predicted values.

There is a fair amount of symmetry through 0, indicating that linearity is present in the model. Therefore, the model is not mispecified.

1. Open the file EXAMSUCCESS
2. Regress EXAMSCORE on CLASSHOURS and interpret your estimate of the coefficient on CLASSHOURS (1 mark)

EXAMSCOREi = β0 + β1CLASSHOURSi + ui

R squared is 0.020, indicating 2% of the variation in examscore is explained by the variation in class hours. Explanatory power of the model is weak.

EXAMSCOREi = 22.730 + 0.121 CLASSHOURSi + ui

For every hour, a student spends in class, their exam score will increase by 0.121.

1. Test the hypothesis that EXAMSCORE is not related to CLASSHOURS (2 marks)

H0: β1 = 0 (EXAMSCORE is not related to CLASSHOURS)

H1: β1 ≠ 0. (EXAMSCORE is related to CLASSHOURS)

T statistic = 3.683, p-value= 0.000 which is less than 0.05.

Therefore, reject H0: β1 = 0 (EXAMSCORE is not related to CLASSHOURS) for significance level of 5%. Examscore and class hours are related.

1. Regress EXAMSCORE on CLASSHOURS and HOMEWORK. Interpret your estimate of the coefficients on CLASSHOURS and HOMEWORK (3 marks)

EXAMSCOREi = β0 + β1CLASSHOURSi +β2HOMEWORKi + ui.

H0: β1 = 0, β2 = 0 (CLASSHOURS & HOMEWORK do not effect EXAMSCORE)

H1: β1 ≠ 0, β2 ≠ 0 (CLASSHOURS & HOMEWORK effect EXAMSCORE)

EXAMSCOREi = 21.801+ 0.083CLASSHOURSi +0.022HOMEWORKi + ui.

Sig> 0.05 for homework and class hours, therefore cannot reject H0, so the variables are not statistically significant for 5%.

For every hour in class a student attends, it will increase their exam score by 0.083. For every 1 percentage point increase of homework a student hands in, it will increase their exam score by 0.022.

1. How much of the variation in EXAMSCORE across students is explained by CLASSHOURS and HOMEWORK? (2 marks)

R2 indicates the percentage variation, and is 0.027. Therefore 2.7% of the variation in Exam Score across students is explained by amount of variation in Class Hours and percentage of homework handed in. 2.7% suggests the explanatory power of the model is fairly weak, and there must be other variables that effect examscore.

1. Why is the estimate of the coefficient on CLASSHOURS different in a) and c)? (2 marks)

The coefficient’s now different as another regressor has been added in part c. Adding the regressor ‘homework’ now allows the estimate of β1 to capture the impact of class hours that does not correlate with percentage of homework handed in. The previous estimate in part a will be overestimated as class hours and percentage of homework handed are positively correlated.

1. Using your results, what advice would you give to students about attending classes and/or working at home? (3 marks)

Advice would be to attend more classes as well as do more work at home, as the regression indicates that an increase in both of these variables will increase the student’s overall exam score.

1. What other variables would you like to include in your regression to help you provide better advice to students? (1 mark)

Other variables which could be included in the variable to provide better advise could be hours of revision, hours’ sleep, and social hours. The more variables included in a regression will increase the explanatory power of the model.

1. Open the file MARKETASSESSMENT .
2. Regress LOGPRICE on RAINFALL and save the predicted values (1 mark)
3. What do the predicted values purport to represent? (3 marks)

Predicted values represents LogPrice^. This is a component of LogPrice which depends upon exogenous variables. The exogenous variable in this case is rainfall. Therefore LogPrice^ is the value of LogPrice which depends on rainfall. It does not include any of LogPrice which depends on LogQuantity.

1. Use the predicted values from a) to gain an estimate of the response of consumer demand to a change in LOGPRICE. (3 marks)

Equation 1: LogQuantityi = β0 + β1Logpricei + ui

Equation 2: LogPricei = β0 + β1LogQuantityi + β2Rainfalli + ui.

Equation 2 ‘purifies’ the data and specifies the reverse causation LogQuantity may have on LogPrice.

The exogenous variable in the model is rainfall, and the endogenous variables are LogPrice and LogQuantity.

The reduced form of the model is LogPricei = β0 + β1 Rainfalli + ui.

Regressing reduced form allows us to regress the equation Logquantityi = β0 + β1^LogPricei + ui.

β1^ = -0.861.

LogQuantityi = 7.873 – 0.861 LogPricei+ ui.

Meaning for a 1% increase in the price, demand will fall by 0.861%.

1. Is the equation that you are trying to estimate in c) identified (3 marks)

The equation LogQuantityi = β0 + β1LOGpricei + ui **is** identified as there is an exogenous variable in the second equation, ‘LogPricei = β0 + β1LOGquantityi + β2Rainfalli + ui’ with a non-zero coefficient which does not appear in equation 1. This exogenous variable is Rainfall.

1. Use the Hausmann test for exogeneity to evaluate whether the responsiveness of consumer demand to price should be estimated using ordinary least squares or two stage least squares (4 marks)

Reduced form: Logpricei = β0 + β1Rainfalli + ui

Saving the unstandardized residuals, add this as another regression into the original equation which becomes:

Logquantityi = β0 + β1LogPricei + β2Res\_1i + ui

Run a Hausmann test, testing for exogeneity.

H0: β2 = 0 (exogenous)

H1: β2 ≠ 0 (endogenous)

Logquantityi = 7.873 – 0.861LogPricei + 0.481Res\_1i + ui

As the p value is 0.282 (>0.05), we cannot reject the hypothesis of exogeneity of LogPrice for significance level 5%, so LogPrice is exogenous, so we should estimate using a one stage test. Using a two stage least squares test will result in unnecessary inefficiency in estimation.

1. Open the file OECDINCOME

In this question we wish to test what percentage of Income Inequality (INEQ), is due to Growth Rate (GR), to unemployment rate (UNEMP), to trade union density (UNDENS), to capital investment (GFCF) based on data from Organisation of Economic Co-operation and Development (OECD).

1. Write down the whole countries regression model for this and estimate the model (1 mark)

Ineqi = β0 + β1GRi + β2Unempi + β3UNDENSi + β4GFCFi + ui

H0: β1=0,β2 = 0,β3 = 0,β4 = 0 (Growth rate, unemployment rate, trade union density, capital investment do not effect inequality)

H1: β1 ≠ 0, β2 ≠ 0, β3 ≠ 0, β4 ≠ 0, (Growth rate, unemployment rate, trade union density, capital investment effect inequality)

R2 = 0.529, 52.9% of the variation in inequality is explained by the variation in the abovementioned variables.

Ineqi = 0.481-0.018GRi + 0.002Unempi + 0.00UNDENSi + 4.076e-14 GFCFi + ui

No multicollinearity as all VIFs < 5.

Reject H0 as All Sigs < 0.05; all variables statistically significant for significance level of 5%.

A 1 percentage point increase in growth rate reduces the Gini coefficient by 0.018.

A 1 percentage point increase in unemployment rate increases Gini coefficient by 0.002.

A 1 percentage point increase in trade union density, will increase the Gini coefficient by 0 to three decimal places.

An increase by 1 to capital investment will increase the Gini coefficient by 4.076e-14.

1. It is believed that the kind of political party in government (right-left) determines the Income Inequality (INEQ) and interacts with the rest of independent variables too. We now introduce a dummy variable for political party in government (POLGOV=1 for left or centre left parties and POLGOV =0 for right wing parties). Test the null hypothesis that after ‘holding all the factors constant (ceteris paribus), POLGOV has no effect on Income Inequality (INEQ). State the alternative hypothesis that POLGOV has an effect on Income Inequality (INEQ). From your regression, what percentage of the variation in Income Inequality (INEQ) is due to the variations in the abovementioned variables? What is the predicted effect of the variables on Income Inequality (INEQ)? What is the p value for their t statistic? Does using a 1%, 5% or 10% significance level make a difference to your decision? (4 marks)

Ineqi = β0 + δ0POLGOVi + β1GRi + β2Unempi + β3UNDENSi + β4GFCFi + ui

H0: δ0 = 0, (ceteris paribus) POLGOV does not affect inequality

H1: δ0 ≠ 0, (ceteris paribus) POLGOV effects inequality

Ineqi = 0.505 -0.006POLGOVi – 0.021GRi + 0.002Unempi + 0.000UNDENSi + 4.12e-14 GFCFi + ui

POLGOVs P-value for t statistic = 0.153 (> 0.05); cannot reject H0: δ0 = 0, POLGOV’s not statistically significant for significance level of 5%, therefore POLGOV does not affect inequality. It’s also not significant for levels 1% and 10% (sig> 0.1 and 0.01).

H0: β1=0,β2 = 0, H0: β3 = 0, H0: β4 = 0 (Growth rate, unemployment rate, trade union density, capital investment effects inequality)

H1: β1 ≠ 0, β2 ≠ 0, β3 ≠ 0, β4 ≠ 0, (Growth rate, unemployment rate, trade union density, capital investment does not affect inequality)

R2 = 0.533, 53.3% of the variation in inequality is explained by the variation in POLGOV, Growth rate, unemployment rate, trade union density, capital investment does not affect inequality.

No evidence of multicollinearity (VIFs < 5).

GR, UNEMP, UDEN, GFCF’s are all statistically significant for 5% and 10% significance level (p-value for t statistics < 0.05 & 0.1). Rejecting H0: β1=0,β2 = 0, H0: β3 = 0, H0: β4 = 0.

GR is statistically insignificant for significance level of 1% as H0: β1=0 rejected (sig> 0.01).

Having a (centre) left government (POLGOV =1) reduces Gini coefficient by 0.006., although it is statistically insignificant.

1 percentage point increase in growth rate reduces Gini coefficient by 0.021.

1 percentage point increase in unemployment rate increases Gini coefficient by 0.002. A 1 percentage point increase in trade union density, will increase the Gini, but is 0.000 to three decimal places.

An increase by 1 to capital investment will increase the Gini coefficient by 4.120e-14.

1. Do countries at years with centre-left governments have a different intercept? (Use the Chow test) (3 marks)

Chow test:

Ineqi = β0 + β1GRi + β2Unempi + β3UNDENSi + β4GFCFi + ui

SSRr (Residual sum of squares, restricted model) = 0.220

Split data into (centre) left and right governments to get SSRur

SSRur (Residual sum of squares, unrestricted model) = 0.073(SSR right wing) + 0.133 (SRR centre/left wing) = 0.206.

= 3.19(2 d.p).

F statistic with 5 parameters and 235 degrees of freedom, for a significance level of 5% is 2.214, and for a significance level of 1% the f statistic is 3.017.

3.19 > 2.214 and 3.017, the hypothesis of a left or centre left government not effecting inequality (H0:δ0 = 0) is rejected. Therefore, a government which is left or left-centre will affect inequality and has different intercept to right wing governments.

1. We wish to see the effect of political party in government on the above regression, in combination with the capital variables and labour variables (POLGOV- GR, POLGOV-UNEMP, POLGOV - UNDENS, POLGOV-GFCF). Can we capture any effect because of political party Do we have a different slope or different intercepts for (centre) left now? Is there any evidence of multicollinearity? (4 marks)

Ineqi= β0 + δ0POLGOVi + β1GRi + δ1POLGOV-GRi+ β2Unempi + δ2 POLGOV-UNEMPi + β3UNDENSi + δ3POLGOV-UNDENSi + β4GFCFi + δ4POLGOV-GFCFi + ui.

H0: β1=0, β2=0, β3=0, β4=0, δ0=0, δ1=0, δ2=0, δ3=0, δ4=0. (variables in equation do not have an effect on inequality).

H1:β1≠0, β2≠0, β3≠0, β4≠0, δ0≠0, δ1≠0, δ2 ≠0,δ3≠0, δ≠04. (variables in equation effect inequality)

R2 = 0.577, 57.7% of the variation in inequality is explained by the variation of variables in the equation.

Ineqi= 0.737-0.411POLGOVi – 0.043GRi + 0.021POLGOV-GRi+ 0.002Unempi -0.001POLGOV-UNEMPi + 0.000UNDENSi + 0.000POLGOV-UNDENSi + 3.537e-14 GFCFi + 0.007POLGOV-GFCFi + ui.

POLGOV, GR, UNEMP, GFCF, POLGOV-GFCF all statistically significant for level of 5%; H0 rejected (sig<0.05). Other variables are not significant; H0 cannot be rejected (sig>0.05).

(Centre) left governments reduce the Gini coefficient by 0.411; different intercept. However, we cannot trust their partial effects because there is strong evidence of multicollinearity. High multicollinearity between POLGOV, POLGOV-GR and POLGOV-GFCF (VIFs very high). Also, multicollinearity with POLGOV-UDEN, POLGOV-UNEMP (VIFs > 5).

A 1 percentage point increase in Growth rate, reduces the Gini coefficient by 0.043.

A 1 percentage point increase in unemployment increases the Gini coefficient by 0.002.

A 1 percentage point increase in trade union density, increases the Gini coefficient by 0 to three decimal places, however this result is statistically insignificant (sig> 0.05, H0: β3= 0 cannot be rejected).

An increase in capital investment by 1, increases the Gini coefficient by 3.537e-14.

An increase by 1 in capital investment within a country with a (centre) left government increases the Gini coefficient by 0.007.

The following interpretations are all statistically insignificant for 5% significance level (sig> 0.05, H0: δ1=0, δ2=0, δ3=0, cannot be rejected):

A 1 percentage point increase in growth rate within a country with a (centre) left government, increases the Gini coefficient by 0.021.

A 1 percentage point increase in unemployment rate within a country with a (centre) left government reduces the Gini coefficient by 0.001.

A 1 percentage point increase in trade union density within a country with a (centre) left government increases the Gini coefficient by 0.000 (3 d.p)

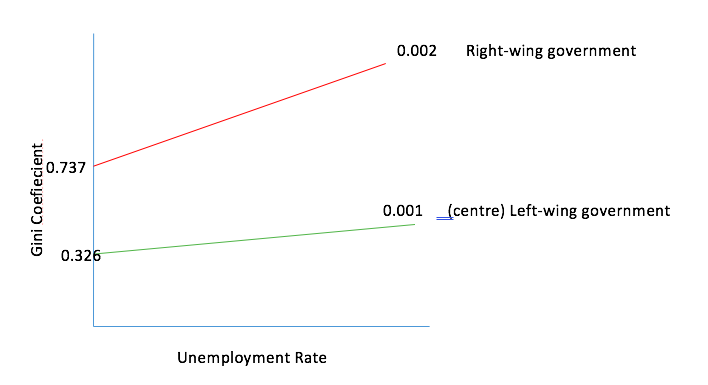
The effect of a (centre) left-wing government, regarding inequality and the unemployment rate can be shown in the graph below.

INEQi = β0 + δ0 POLGOVi + β1UNEMPi + δ1POLGOV-UNEMPi + ui

(centre) Left-wing government (POLGOV = 1): INEQi = 0.737 – 0.411+ 0.002 UNEMPi – 0.001 UNEMPi + ui

INEQi = 0.326+ 0.001 UNEMPi  + ui

Right-wing government (POLGOV = 0): INEQi = 0.737 + 0.002 UNEMPi  + ui



From the graph, we can see Inequality is higher when a country has a Right-wing government as this intercept is higher. It is also evident that the Gini coefficient increases due to unemployment at a faster rate in a right-wing country than in ones with (centre) left-wing governments.

1. Now we want to see what is the probability of a country to have a (centre) left government (POLGOV =1) depending on the abovementioned variables (Income Inequality (INEQ), is due to Growth Rate (GR), to unemployment rate (UNEMP), to trade union density (UNDENS) and to capital investment (GFCF))? Use the Linear Probability Model (Ordinary Least Squares –OLS) and the LOGIT Model (Maximum Likelihood Estimations –MLE) and interpret the results. Compare the two models (8 marks)

**Linear Probability model**

P(POLGOV=1/xi) = β0 +β1INEQi + β2GRi + β3UNEMPi + β4UNDENSi +β5 GFCFi + ui

H0: β1=0, β2=0, β3=0 β4=0 β5=0 (Inequality, growth rate, unemployment rate, trade union density, capital investment does not affect the probability of a country having a (centre) left government)

Ha: β1≠0, β2≠0, β3≠0 β4≠0 β5≠0 (Inequality, growth rate, unemployment rate, trade union density, capital investment affects the probability of a country having a (centre) left government)

R2 = 0.064, 6.4% of the variation the probability of a country having a (centre) left party is due to the above-mentioned variables.

P(POLGOV=1/xi) = 4.850– 1.493 INEQi – 0.389 GRi – 0.002 UNEMPi + 0.005 UNDENSi + 1.358e-13 GFCFi+ ui

Only Growth rate and trade union density are statistically significant for 5% significance level; H0 rejected (sig<0.05).

H0 cannot be rejected for 5% significance level for the other variables (sig>0.05).

No evidence of multicollinearity (VIFs < 5)

1 percentage point increase in inequality, reduces the probability of a (centre)left wing government by 149.3%. However, this result is statistically insignificant (sig> 0.05). This is a disadvantage with the LPM, as probabilities are between 1 and 0, therefore this effect does not make sense.

1 percentage point increase in growth rate, reduces the probability of a (centre)left wing government by 38.9%

1 percentage point increase in unemployment rate reduces the probability of a (centre)left wing government by 0.2%. However, this result is statistically insignificant (sig> 0.05).

1 percentage point increase in trade union density increases the probability of a (centre)left wing government by 0.5%.

An increase in Capital investment by 1, increases the probability of a (centre)left wing government by 1.358e-13. However, this result is statistically insignificant (sig> 0.05).

**Logit**

P(POLGOV=1/xi) = G(β0 +β1INEQi + β2GRi + β3UNEMPi + β4UNDENSi +β5 GFCFi + ui)= G(z)

G(z) =

H0: β1=0, β2=0, β3=0 β4=0 β5=0

H1: β1≠0, β2≠0, β3≠0 β4≠0 β5≠0

Block 0:

No predicted variables that are in Right-wing countries.

Assuming every country has a (centre) left-wing government, we get 52.1% classification accuracy.

Sig > 0.05, reject the null hypothesis that there is an equal number of data within sampling variability in the Right-wing and (centre) left-wing.

With no explanatory variables, the intercept is 0.083.

Exp(B) = 1.087; 108.7% likelihood of having a (centre) left-wing government.

Variables not in the equation table indicates the variables GINI, GR, UNEMP and GFCF are not statistically significant for level of 5% (reject H0 as sig> 0.05). UDEN is statistically significant.

Block 1:

Sig< 0.05; reject hypothesis that there is some predictive capacity in the equation. Evidence that the variables can predict the probability of a country having a (centre) left or right-wing government.

Nagelkerke R2= 0.086; roughly, 8.6% of the variation in the probability of a country having a (centre) left-wing government is explained by the variation in inequality, growth rate, unemployment rate, trade union density and capital investment.

Hosmer and Lemeshow test shows that Chi-square is not statistically significant (sig> 0.05); no evidence of model misspecification, the productive capacity of the model is quite high.

Classification table now shows :

Block 0 overall percentage prediction = 52.1%

Block 1 overall percentage prediction = 55.8%

The productive capacity of the model compared to the No Model has increased by 3.7%.

An increase in the Gini coefficient by 1, reduces the probability of a (centre) left-wing government by 99.8%, but the result is not statistically significant (sig>0.05).

An increase in the growth rate by 1 percentage point reduces the probability of a (centre) left-wing government by 81%.

An increase in the unemployment rate by 1 percentage point reduces the probability of a (centre) left-wing government by 0.8%, but the result is not statistically significant (sig>0.05).

An increase in trade union density by 1 percentage point increases the probability of a (centre) left-wing government by 2.4%.

An increase in capital investment by 1 increases the probability of a (centre) left-wing government by 0.000% (3 d.p), but the result is not statistically significant (sig>0.05).

Comparing the two models (Logit and LPM), the logit model has higher explanatory power as it has a higher R2, however, the R2 in the Logit model is a pseudo R2 and isn’t the best way to calculate explanatory power.

The Logit model removes probabilities which do not make sense such as ones with over 100%. Shown by an increase in the Gini coefficient by 1 reduces the probability of a (centre) left government by 149.3% in the LPM. Whereas in logit, an increase in the Gini coefficient by 1, reduces the probability of a (centre) left-wing government by 99.8%. Making the Logit model easier to interpret and make sense of.

1. Open the file CONSINCOME.
2. Is the variable CONSUMPTION stationary or non-stationary? If it is

non-stationary what form of non-stationarity does it exhibit? (6 marks)

ChangeConsumptiont = β0 + β1 LagConsumptiont + β2 Timet + ut.

H0: β2 = 0 (No deterministic non-stationarity)

H1: β2 ≠ 0 (deterministic non-stationarity)

Sig > 0.05, do not reject H0 for 5% significance level. No deterministic non-stationarity exhibited in Consumption.

Perform Dickey-fuller test for unit root:

H0: β1 = 0 (random non-stationarity/ unit root)

H1: β1 ≠ 0 (no random non-stationarity/ no unit root)

T statistic = -2.394, when looking on table 18.2, the statistic gives a p value greater than 5%.

Therefore, we cannot reject H0 for 5% significance level; Consumption has a unit root and random non-stationarity.

1. Is the variable INCOME stationary or non-stationary? If it is

non-stationary what form of non-stationarity does it exhibit? (6 marks)

ChangeIncomet = β0 + β1 LagIncomet + β2 Timet + ut.

H0: β2 = 0 (No deterministic non-stationarity)

H1: β2 ≠ 0 (deterministic non-stationarity)

Sig > 0.05; cannot reject H0 for 5% significance level, Income does not exhibit deterministic non-stationarity.

Dickey-fuller test for unit root:

H0: β1 = 0 (random non-stationarity/ unit root)

H1: β1 ≠ 0 (no random non-stationarity/ no unit root)

T statistic = -1.179, on table 18.2 will give a p-value greater than 5%. Therefore, cannot reject H0; income has unit root and exhibits random non-stationarity.

1. Conduct a regression to test the relationship between CONSUMPTION and INCOME and justify your choice of technique. (8 marks)

Running a normal regression of consumption on income results in r2 0.280.

However, as consumption and income are both non-stationary time series, we must test to see if this relationship between consumption and income is truly meaningful, which is called cointegration. To test for cointegration we test for the presence of a unit root when regressing the change in residuals on the lag residuals. If there is no unit root, it suggests that cointegration is exhibited.

H0: β1 = 0 (unit-root in residuals, no cointegration)

H1: β1 ≠ 0 (no unit-root in residuals, cointegration)

T-statistic = -2.187, on table 18.4, gives p value greater than 5%; cannot reject H0. Meaning there is a unit root in the residuals so there is no cointegration.

This indicates that the regression of the two non-stationary time series was a spurious regression and that the two variables do not have a meaningful relationship.