**ECON 3320 Applied Econometrics**

|  |  |
| --- | --- |
| Critical Values | Z0.10 Z0.05 Z0.025 Z0.01  |
| **One tail values** |  1.295 1.645 1.96 2.33 |

**1. PART A: 50 points**

The researchers are interested in modeling the birth weight of children. They form the following model:

where the variables are defined as:

* **BWGHT**: child birth weight (in grams)
* **CIGS**: average number of cigarettes smoked by the mother per day
* **FEDUC**: years of education of the **father**
* **MEDUC**: years of education of the **mother**
* **MAGE**: age of the mother (in years)
* **MALE**: 1 if the child is a male, 0 otherwise
* **NPVIS**: total number of prenatal visits

By using an Ordinary Least Square Method (OLS) the following regression model is estimated:

|  |
| --- |
| **TABLE 1**Dependent Variable: **LOG(BWGHT)**  |
| Included observations: 1625  |
| **Variable**  | **Coefficient** | **Std. Error**  |
| **Constant**  | 7.711289 | 0.138997  |
| **CIGS**  | -0.002397 | 0.001184  |
| **FEDUC**  | 0.003066 | 0.002673  |
| **MEDUC**  | -0.002564 | 0.002975  |
| **MAGE**  | 0.021644 | 0.009502  |
| **MAGE^2**  | -0.000345 | 0.000159  |
| **MALE**  | 0.025137 | 0.009748  |
| **NPVIS**  | 0.004978 | 0.001320  |
| R-squared 0.021706DW 1.56 |  |

answer the following questions based on the given estimated model result in Table 1:

1. We are interested to test whether mother’s smoking has a ***negative effect*** on child birth weight. Test this hypothesis using a 5% significance level.
2. What is the effect of **10more**prenatal visits on the child birth weight?
3. Interpret the coefficient estimate of the **MALE** dummy variable .
4. Based on the DW statistic in **Table 1**, decide whether you reject or not the null of no serial correlation of the error.(From DW critical table: dL = 1.697 dU = 1.841; HINT: create the boundries for positive , negative and inconclusive decision areas)
5. If there is a perfectly positive autocorrelation in residuals in first order, what is the remedy to eliminate the first order autocorrelation? Explain it briefly. You can use an example.
6. We estimate the regression model for the squared residuals (RESID^2)(in other words, error terms square: **(**shown in the **Table 2 (below)**. How would you implement a test for heteroskedasticity based on this regression results? State the null and alternative hypotheses, and calculate the χ2 test statistics as χ2=T\*R2. (use a 5% significance level in your decision).

Hint: White-Test is a heteroscadasticity test which takes error term squares and regresses it on the independent variables, their squares and their interactions. You need to test if all coefficients are **jointly equal to zero or not**. It is basically an F-test for the significancy for the whole model. If F-test is not given, use χ2=T\*R2 where T is the number of observations and R2 of the regression. This will give you the χ2 test statistics to compare it with χ2(k=14) critical value, which is 23.7 from the χ2 critical table.

So compare 23.7 and χ2=T\*R2 =1625\***0.026298=???**

|  |
| --- |
| **TABLE 2**Dependent Variable: **RESID^2()** |
| Variable  | Coefficient  | Std. Error  |
| **Constant**  | 1.797312  | 2.044727  |
| **CIGS**  | 0.029079  | 0.032139  |
| **CIGS^2**  | 7.82E-05  | 0.000127  |
| **CIGS\*MAGE**  | -0.002281  | 0.002237  |
| **CIGS\*(MAGE^2)**  | 3.73E-05  | 3.81E-05  |
| **CIGS\*MALE**  | 0.000682  | 0.002298  |
| **CIGS\*NPVIS**  | 0.000168  | 0.000205  |
| **MAGE**  | -0.266978  | 0.288828  |
| **MAGE^2**  | 0.014901  | 0.015135  |
| **MAGE\*(MAGE^2)**  | -0.000337  | 0.000346  |
| **MAGE\*MALE**  | 0.003528  | 0.018641  |
| **MAGE\*NPVIS**  | -0.004875  | 0.002625  |
| **(MAGE^2)^2**  | 2.62E-06  | 2.91E-06  |
| **(MAGE^2)\*MALE**  | -3.66E-05  | 0.000313  |
| **(MAGE^2)\*NPVIS**  | 8.25E-05  | 4.41E-05  |
| **MALE**  | -0.067026  | 0.274525  |
| **MALE\*NPVIS**  | -0.000336  | 0.002541  |
| **NPVIS**  | 0.050867  | 0.038389  |
| **NPVIS^2**  | 0.000477  | 0.000115  |
| R-squared **0.026298** |  |

1. Based on the results of the White tests, do you think that OLS is the appropriate estimator? What goes wrong if we use OLS and the errors are heteroskedastic?
2. Discuss the problem of heteroskedasticity. what is the nature of the problem. Why do you see a heteroscadasticity (non-constant variance) problem? Give an example.

**PART B: (50 points):**

1. You are given the below model for the stock returns:

(1)

TABLE 1:

Dependent Variable: LOG(MICROSOFT)

Method: Least Squares

Date: 04/28/14 Time: 17:22

Sample: 2002M01 2007M04

Included observations: 64

Variable Coefficient Std. Error t-Statistic Prob.

C -0.930703 0.338686 -2.747978 0.0078

LOG(SANDP) 0.585776 0.048161 12.16285 0.0000

R-squared 0.704670 Mean dependent var 3.187797

Adjusted R-squared 0.699907 S.D. dependent var 0.102948

S.E. of regression 0.056396 Akaike info criterion -2.882103

Sum squared resid 0.197188 Schwarz criterion -2.814638

Log likelihood 94.22731 Hannan-Quinn criter. -2.855525

F-statistic 147.9349 Durbin-Watson stat 0.902327

Prob(F-statistic) 0.000000

1. Interpret the R2 in Table 2.

2. Test if the model as a whole statistically significant or not. Write the null hypothesis. Be clear which test you use. Write your test statistic, critical value at 5% significance level. Write down what is p value of the test statistic. Compare it with your significance level (5%).

4. The below graph is taken from the regression estimation. Do you suspect any autocorrelation in the residual series visually. Explain. If there is no autocorrelation, what type of visual graph you expect to see? Draw it hypothetically.



5. We want to test *serial correlation*by using *the LM Serial correlation* test for **2** lags of the residuals. The estimation result is given in Table 3.

a.Specify the null hypothesis of the test,

b. use F-statistic to decide on the null hypothesis at 5 % significance level.

c. calculate the χ2 test statistic and decide whether you reject at 5%. (Hint χ2 test statistic= T\*R2 and distributed as χ2 (m), m is number lags for serial correlation.χ2 (2)= 5.99 from the critical table)

**TABLE 3:Breusch-Godfrey Serial Correlation LM Test**

F-statistic 10.75610 Prob. F(2,60) 0.0001

Obs\*R-squared 16.89049 Prob. Chi-Square(2) 0.0002

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 04/28/14 Time: 17:33

Sample: 2002M01 2007M04

Included observations: 64

Presample missing value lagged residuals set to zero.

Variable Coefficient Std. Error t-Statistic Prob.

C -0.037668 0.295718 -0.127378 0.8991

LOG(SANDP) 0.005409 0.042052 0.128633 0.8981

RESID(-1) 0.553291 0.129711 4.265562 0.0001

RESID(-2) -0.079426 0.129790 -0.611959 0.5429

**R-squared 0.263914** Mean dependent var -3.38E-17

Adjusted R-squared 0.227110 S.D. dependent var 0.055946

S.E. of regression 0.049185 Akaike info criterion -3.126012

Sum squared resid 0.145147 Schwarz criterion -2.991081

Log likelihood 104.0324 Hannan-Quinn criter. -3.072856

F-statistic 7.170737 Durbin-Watson stat 1.896615

Prob(F-statistic) 0.000341

d. What is your decision for the null hypothesis? Did you reject or not? Is there an autocorrelation in the residuals in second order?

6. The below regression model is given for the excess return of Microsoft return and excess return of the market:

 (2)

where:

R denotes a specific stock return at time t

RF denotes risk free interest rate at time t, generally US Treasury Bill interest rate.

RM denotes the market return at time t. Generally S&P 500 Index, or Dow Jones Index as a proxy for market return (all stocks)

**TABLE 4:**

**Dependent Variable: ERMICROSOFT**

Method: Least Squares

Date: 04/28/14 Time: 17:41

Sample (adjusted): 2002M02 2007M04

Included observations: 63 after adjustments

Variable Coefficient Std. Error t-Statistic Prob.

C -0.320282 0.648445 -0.493923 0.6231

ERSANDP 1.070772 0.182660 5.862098 0.0000

R-squared 0.360347 Mean dependent var 0.121478

Adjusted R-squared 0.349861 S.D. dependent var 6.339973

S.E. of regression 5.111997 Akaike info criterion 6.132288

Sum squared resid 1594.083 Schwarz criterion 6.200324

Log likelihood -191.1671 Hannan-Quinn criter. 6.159047

F-statistic 34.36420 **Durbin-Watson stat 2.212775**

The claim is the excess return of Microsoft return is less riskier than the market return. Test if slope coefficient is equal to 1 in favor of less than 1.

7. Based on the estimation of the regression (2), we have tested Serial Correlation for lag 2. The result is given in Table 5. Based on the estimation, what do you conclude on the serial correlation of the regression? Is there any autocorrelation in regression (2)

TABLE 5: **Breusch-Godfrey Serial Correlation LM Test:**

**F-statistic 0.906154 Prob. F(2,59) 0.4096**

**Obs\*R-squared 1.877505 Prob. Chi-Square(2) 0.3911**

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 04/28/14 Time: 17:46

Sample: 2002M02 2007M04

Included observations: 63

Presample missing value lagged residuals set to zero.

Variable Coefficient Std. Error t-Statistic Prob.

C -0.016228 0.649558 -0.024983 0.9802

ERSANDP 0.018453 0.183465 0.100583 0.9202

RESID(-1) -0.137362 0.129684 -1.059210 0.2938

RESID(-2) -0.123842 0.129873 -0.953560 0.3442

R-squared 0.029802 Mean dependent var 8.46E-17

Adjusted R-squared -0.019530 S.D. dependent var 5.070603

S.E. of regression 5.119879 Akaike info criterion 6.165526

Sum squared resid 1546.577 Schwarz criterion 6.301598

Log likelihood -190.2141 Hannan-Quinn criter. 6.219044

F-statistic 0.604103 Durbin-Watson stat 1.979831

Prob(F-statistic) 0.614932