Examination Period 3: 2016/17

**ECNM01217N**

<table>
<thead>
<tr>
<th>Module Title</th>
<th>Econometrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Seven</td>
</tr>
<tr>
<td>Time Allowed</td>
<td>Two hours</td>
</tr>
</tbody>
</table>

Instructions to students:

- Enter your student number **not** your name on all answer books.
- Answer any **three** questions.
- All questions are equally weighted. Where a question has more than one part the division of marks is stated.
- Begin each answer in a separate booklet; label each booklet clearly with the number of the question you are answering.
- The use of a non-programmable calculator is permitted.
- Statistical tables are provided in a separate document.

<table>
<thead>
<tr>
<th>No. of Pages</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Questions</td>
<td>5</td>
</tr>
</tbody>
</table>
Answer three out of five questions

1. A 2013 study in the USA investigated the relationship between employees’ wages, gender, race, union membership, education and work experience. Equation 1 is the least squares estimated model (standard errors in parentheses):

   **Equation one:**
   \[
   Wage_i = -7.183 - 3.08D_{1i} - 1.556D_{2i} + 1.116D_{3i} + 1.371Education_i + 0.166Experience_i
   \]
   \[\text{se} = (1.015) (0.364) (0.509) (0.508) (0.066) (0.016)\]

   \[R^2=0.234, \ n=1,289\]

   Where Wage is measured in thousands of dollars; \(D_{1i}=1\) for Female, 0 for male; \(D_{2i}=1\) for Non-white, 0 for white; \(D_{3i}=1\) if union member, 0 for non-union member; education is number of years of education and experience is number of years of work experience.

   a. Interpret the regression results from equation one. Conduct appropriate t and F tests for the significance of all independent variables in the model at 95% level of significance.

   **(10 marks)**

   b. Explain what is heteroskedasticity and outline the consequences of heteroskedasticity.

   **(5 marks)**

   c. How can we detect heteroskedasticity?

   **(5 marks)**

   d. What are the remedies for heteroskedasticity?

   **(5 marks)**

   e. Consider the below test results for heteroskedasticity in equation one. Interpret these results.

   **(8 marks)**

   **(33 marks total)**
2. A study in the UK between 1986-2015 investigated the demand of ice cream based on the following model (standard errors in parentheses):

**Equation one:**
\[ \hat{Q}_t = 0.197 + 0.003I_t - 1.044P_t + 0.003F_t \]
\[ se= (0.270) (0.001) (0.834) (0.001) \]

R\(^2\)= 0.719, DW= 1.021

Where Q is per capita consumption of ice cream in pints, I is weekly family income in British Pounds, P is price per pint in British Pounds and F is mean temperature in Fahrenheit.

a. What are a priori expectations on the anticipated signs of the regression coefficients? Interpret the regression results from equation one. Are all the independent variables significant in explaining the dependent variable?  (8 marks)

b. Explain what autocorrelation is. Use the results from equation one to test for autocorrelation/serial correlation and interpret the findings.  (10 marks)

c. Consider the below test for higher order autocorrelation/serial correlation in equation two (standard errors in parentheses). Interpret these results.

**Equation two:**
\[ \hat{Q}_t = 0.029 + 0.004I_t - 0.708P_t + 0.003F_t + 0.099Q_{t-1} \]
\[ se= (0.270) (0.001) (0.310) (0.001) (0.048) \]

R\(^2\)= 0.765, DW = 1.176

**Results table:**

<table>
<thead>
<tr>
<th>Breusch-Godfrey Serial Correlation LM Test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>Observations*R-squared</td>
</tr>
</tbody>
</table>

(10 marks)

d. What are the remedies for autocorrelation?  (5 marks)
3. We consider the following demand function for chicken in the United States between 1994-2016:

**Equation one:**

\[
\ln Y_t = \beta_0 + \beta_1 \ln I_t + \beta_2 \ln P_{\text{chicken}}_t + \beta_3 \ln P_{\text{pork}}_t + \beta_4 \ln P_{\text{beef}}_t + u_t
\]

Where \( Y \) is per capita consumption of chicken in pounds (lb), \( I \) is real disposable per capita income in US dollars, \( P_{\text{chicken}} \) is the real retail price of chicken per pound (lb) in dollars, \( P_{\text{pork}} \) is the real retail price of pork per pound (lb) in dollars, \( P_{\text{beef}} \) is the real retail price of beef per pound (lb) in dollars and \( \ln \) denotes the natural logarithm.

The estimated model from **equation one** is as follows (standard errors in parentheses):

**Equation two:**

\[
\ln \hat{Y}_t = 2.189 + 0.343 \ln I_t - 0.505 \ln P_{\text{chicken}}_t + 0.149 \ln P_{\text{pork}}_t + 0.091 \ln P_{\text{beef}}_t
\]

\[\text{se}=(0.156) \quad (0.083) \quad (0.111) \quad (0.010) \quad (0.101)\]

\( R^2 = 0.982, \text{RSS} = 0.014 \)

**a.** What are a priori expectations on the anticipated signs of the regression coefficients? Interpret the regression results from **equation two**.

(8 marks)

**b.** Conduct appropriate t and F tests for the significance of all independent variables in the model at 95% level of significance.

(10 marks)

Question three continues overleaf
c. Suppose that a researcher supports that the demand of chicken is not affected by the price of pork and the price of beef. **Equation one** then becomes:

**Equation three:**

\[ \ln Y_t = \beta_0 + \beta_1 \ln I_t + \beta_2 \ln P_{chicken} + u_t \]

and the estimated model is as follows (standard error in parentheses):

**Equation four:**

\[ \ln \hat{Y}_t = 2.033 + 0.452 \ln I_t - 0.377 \ln P_{chicken} \]

se = (0.116) (0.025) (0.064)

\[ R^2 = 0.98, \quad RSS = 0.015 \]

Use the above results to test the hypothesis that the demand of chicken is not affected by the price of pork and the price of beef.

(15 marks)

(33 marks total)

4.

a. What is meant by stationarity? What is an integrated variable?

(8 marks)

b. Explain the relationship between stationarity and a spurious regression; why does the OLS estimators give such problematic results?

(8 marks)

c. How would you assess the stationarity of a variable?

(17 marks)

(33 marks total)

---

**Question five continues overleaf**
We consider the following Logit model to explain how students’ final grades on Macroeconomics depend on grades on an initial Macroeconomics test, a new teaching method (NTM) and their entering grade point average (GPA).

**Equation one:**
\[ L_i = \left( \frac{p_i}{1-p_i} \right) = \beta_1 + \beta_2 InitialTest_i + \beta_3 NTM_i + \beta_4 GPA + u_i \]

The estimated model is as follows (standard errors in parentheses):

**Equation two:**
\[ L_i = -13.021 + 0.095 InitialTest_i + 2.379 NTM_i + 2.826 GPA \]
se = (4.931) (0.142) (1.065) (1.263)
McfFadden R²=0.374, n=16

Where the final grade=1, if final grade is A and 0 if final grade is B or C; InitialTest is the score on an examination given at the beginning of the term to test the knowledge of macroeconomics; NTM=1, if the new teaching method is used, 0 otherwise; and GPA is the entering grade point average.

a. What are the problems with linear probability models? Why is there a need for a logit or probit model?  
(10 marks)

b. What are the characteristics of the logit model? What is the difference between the logit and the probit model?  
(10 marks)

c. What are a priori expectations on the anticipated signs of the regression coefficients? Interpret the regression results from equation two. 
(13 marks)

(33 marks total)