Examination Period 3: 2018/19

**ECNM01219N**

<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 **four** out of **six** questions.
- Begin each question in a separate answer book; label each answer book clearly with the number of the question you are answering.
- The use of a calculator **is** permitted.

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

1. Consider production function of the form $Q = f(H, K)$, where $Q$ is the output measure and $H$ and $K$ are hours worked and gross capital stock, respectively. Based on 33 observations we obtain the following results:

$$\log(Q) = 0.129 + 0.448 \log(K) + 0.559 \log(H)$$

$(se)$ $(0.546)$ $(0.704)$ $(0.816)$ $R^2 = 0.886$

a. Interpret the regression results. (8 marks)

b. What is the output elasticity of hours worked? (5 marks)

c. Verify that the coefficients of $\log(H)$ and $\log(K)$ are statistically insignificant at the 5% level. (5 marks)

d. What might account for the insignificance of $\log(K)$ and $\log(H)$ if you are told that the correlation coefficient between $\log(H)$ and $\log(K)$ is 0.980? (7 marks)

Total: 25 marks
2. The following model is a simplified version of the multiple regression model used by Biddle and Hamermesh (1990) to study the tradeoff between time spent on sleeping and working and to look at other factors affecting sleep:

\[
Sleep = \beta_0 + \beta_1 totwork + \beta_2 Educ + \beta_3 age + \beta_4 age^2 + \beta_5 male + u
\]

The variable \(sleep\) is total minutes per week spent sleeping at night, \(totwork\) is total weekly minutes spent working, \(Educ\) and \(age\) are measured in years, and \(male\) is a gender dummy.

Using the data in SLEEP75 (From Biddle and Hamermesh, 1990), we obtain the estimated equation:

\[
\hat{Sleep} = 3,840.83 - 0.163 totwork - 11.71 Educ - 8.70 age \\
(235.11) (0.018) (5.86) (11.21) \\
+ 0.128 age^2 + 87.75 male \\
(0.134) (34.33)
\]

\(n = 706, R^2 = 0.123, \hat{R}^2 = 0.117.\)

a. All other factors being equal, is there evidence that men sleep more than women? How strong is the evidence?  
(5 marks)

b. Is there a statistically significant tradeoff between working and sleeping? What is the estimated tradeoff?  
(8 marks)

c. What other regression do you need to run to test the null hypothesis that, holding other factors fixed, age has no effect on sleeping?  
(5 marks)

d. Explain the role of dummy variables in regression models.  
(7 marks)

Total: 25 marks
3. Consider the regression of consumption of food on a constant and the disposable income. The econometrician suspects that the residuals of this model might be autocorrelated.

   a. She obtained the following Ljung & Box Q-statistics and the corresponding p-values for the estimated residuals. Perform the test with 1% significance level. What is your conclusion?

<table>
<thead>
<tr>
<th>Up to Lag</th>
<th>Q-Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.562</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>12.896</td>
<td>0.002</td>
</tr>
</tbody>
</table>

   (9 marks)

   b. She is also considering the Breusch-Godfrey Serial Correlation LM test with lagged residuals up to order 4. Explain the steps involved in testing. She obtained a value for the LM test statistic of 11.67489. Perform the test with 5% significance level. What is your conclusion?

   (8 marks)

   c. In addition to autocorrelation, the econometrician also likes to check if the residuals are normally distributed. The Jarque-Bera test statistic is 0.066322 with the corresponding p-value of 0.97. Perform the test. What is your conclusion?

   (8 marks)

   Total: 25 marks

4. We are interested in the relationship between rice production, inputs of labour and fertilizer using data on 44 farms.

   \[ RICE_i = B_1 + B_2LABOR_i + B_3FERT_i + \epsilon_i \]

   a. We observe the estimated OLS residuals from the above regression increase in magnitude when plotted against the area planted (measured in acres). Let’s name the variable - the area planted ACRES. We regress \( \hat{\epsilon}_i^2 \) on ACRES and a constant. This regression gives a R² of 0.2068. What can we conclude about heteroskedasticity considering the Breusch-Pagan test with the 1% significance level?

   (9 marks)

   b. We instead estimate the model:

   \[ \frac{RICE_i}{ACRES_i} = a + B_1 \left( \frac{1}{ACRES_i} \right) + \frac{LABOR_i}{ACRES_i} + B_3 \frac{FERT_i}{ACRES_i} + \epsilon_i \]

   What is the implicit assumption about the heteroscedasticity pattern?

   (8 marks)

   c. Many economists would omit \( \frac{1}{ACRES} \) from the equation in b. What argument can you propose that would make this defensible?

   (8 marks)

   Total: 25 marks
5. From the household budget survey of 1980 of the Dutch Central Bureau of Statistics, J. S. Cramer obtained the following logit model based on a sample of 2820 households. (The results given here are based on the method of maximum likelihood and are after the third iteration.)** The purpose of the logit model was to determine car ownership as a function of (logarithm of) income. Car ownership was a binary variable: \( Y = 1 \) if a household owns a car, zero otherwise.

\[
\hat{L}_i = -2.77231 + 0.347582 \ln \text{Income}
\]

\[
t = (-3.35) \quad (4.05)
\]

\[
\chi^2 (1 \text{ df}) = 16.681 \ (p \text{ value} = 0.0000)
\]

where \( \hat{L}_i \) estimated logit and where \( \ln \text{Income} \) is the logarithm of income. The \( \chi^2 \) measures the goodness of fit of the model.

a. Interpret the estimated logit model. \( (7 \text{ marks}) \)

b. From the estimated logit model, how would you obtain the expression for the probability of car ownership? \( (7 \text{ marks}) \)

c. What is the probability that a household with an income of 20,000 will own a car? And at an income level of 25,000? What is the rate of change of probability at the income level of 20,000? \( (6 \text{ marks}) \)

d. Comment on the statistical significance of the estimated logit model. \( (5 \text{ marks}) \)

Total: 25 marks

6. Discuss Best Liner Unbiased Estimators (BLUE) and explain what are the cause and consequences of multicollinearity. \( \text{Total: 25 marks} \)

---

End of Paper