Examination Period 3: 2017/18

ECNM01218N

Module Title  Econometrics  
Level  Seven  
Time Allowed  Two hours

Instructions to students:
- Enter your student number **not** your name on all answer books.
- Answer **three** out of **five** questions.
- Begin each answer in a separate answer book; label each answer book clearly with the number of the question you are answering.
- The use of a non-programmable calculator is permitted.
- Normal distribution table and statistical tables will be provided.

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

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

\( R^2 = 0.886 \)

\( \text{se} \) \( (0.546) \) \( (0.704) \) \( (0.816) \)

a. Interpret the regression results. \hspace{1cm} (10 marks)

b. What is the output elasticity of hours worked? \hspace{1cm} (8 marks)

c. Verify that the coefficients of \( \log(H) \) and \( \log(K) \) are statistically insignificant at the 5% level. \hspace{1cm} (7 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? \hspace{1cm} (9 marks)

Total: 34 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:

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

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:

\[
\text{Sleep} = 3,840.83 - 0.163 \text{totwork} - 11.71 \text{Educ} - 8.70 \text{age}
+ 0.128 \text{age}^2 + 87.75 \text{male}
\]

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

Question 2 continues overleaf
a. All other factors being equal, is there evidence that men sleep more than women? How strong is the evidence?  

(8 marks)

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

(9 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?  

(8 marks)

d. Explain the role of dummy variables in regression models.  

(8 marks)

Total: 33 marks

3. In studying the movement in the production workers’ share in the value added (i.e., labor’s share), the following models were considered by Gujarati:

\[ Model A: Y_t = \beta_0 + \beta_1 t + u_t \]
\[ Model B: Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + u_t \]

Where \( Y \) is labor’s share and \( t \) is time. Based on annual data for 1949 – 1964, the following results were obtained for the primary metal industry:

\[ Model A: \hat{Y}_t = 0.4529 - 0.0041 t \quad R^2 = 0.5284 \quad d = 0.8252 \]
\[ (-3.9608) \]

\[ Model B: \hat{Y}_t = 0.4786 - 0.0127 t + 0.0005 t^2 \quad R^2 = 0.6629 \quad d = 1.82 \]
\[ (-3.2724) \quad (2.7777) \]

a. Is there serial correlation in model A? In model B?  

(8 marks)

b. What accounts for the serial correlation?  

(8 marks)

c. How would you distinguish between “pure” autocorrelation and specification bias?  

(8 marks)

d. Consequences of autocorrelation can be very serious. Explain one possible remedial measure that can redress the problem of autocorrelation.  

(9 marks)

Total: 33 marks
4. 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. 

b. From the estimated logit model, how would you obtain the expression for the probability of car ownership?

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?

d. Comment on the statistical significance of the estimated logit model.

Total: 33 marks
5. For pedagogic purposes Hanushek and Jackson estimate the following model:

\[ C_t = \beta_1 + \beta_2 \text{GNP}_t + \beta_3 D_t + u_t \quad (1) \]

Where \( C_t \) is aggregate private consumption expenditure in year \( t \), \( \text{GNP}_t \) is gross national product in year \( t \), and \( D_t \) is national defence expenditures in year \( t \), the objective of the analysis being to study the effect of defence expenditures on other expenditures in the economy.

Postulating that \( \sigma^2 = \sigma^2 (\text{GNP}_t)^2 \), they transform (1) and estimate

\[ \frac{C_t}{\text{GNP}_t} = \beta_1 (1/\text{GNP}_t) + \beta_2 + \beta_3 (D_t/\text{GNP}_t) + u_t/\text{GNP}_t \quad (2) \]

The empirical results based on the data for 1946–1975 were as follows (standard errors in the parentheses):

\[
\begin{align*}
\hat{C}_t &= 26.19 + 0.6248 \text{GNP}_t - 0.4398 D_t \\
&= (2.73) \quad (0.0060) \quad (0.0736) \\
R^2 &= 0.999 \\
\hat{\text{GNP}}_t &= 25.92 (1/\text{GNP}_t) + 0.6246 - 0.4315 (D_t/\text{GNP}_t) \\
&= (2.22) \quad (0.0068) \quad (0.0597) \\
R^2 &= 0.875
\end{align*}
\]

a. What assumption is made by the authors about the nature of heteroscedasticity? Can you justify it? (12 marks)

b. Compare the results of the two regressions. Has the transformation of the original model improved the results, that is, reduced the estimated standard errors? Why or why not? (10 marks)

c. Can you compare the two \( R^2 \) values? Why or why not? (11 marks)

Total: 33 marks