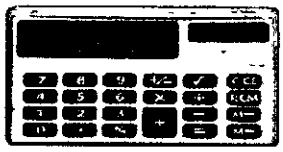
UNIVERSITY EXAMINATIONS



UNIVERSITEITSEKSAMENS

UNISA university of south africa

ECS3706

May/June 2016

ECONOMETRICS

Duration 2 Hours

100 Marks

EXAMINERS

FIRST	MR MJ KHUMALO
SECOND	MR K LELAKA
EXTERNAL	MR SM MILLARD

Use of a non-programmable pocket calculator is permissible.

Closed book examination

This examination question paper remains the property of the University of South Africa and may not be removed from the examination venue.

This paper consists of 26 pages, including a formulae sheet (p18), 3 pages of statistical tables (pp 19 to 21), 5 pages for rough work (pp 22 to 26) plus the special front page

[TURN OVER]

This paper consists of two sections

Section A: Answer all 4 questions which together count 60 marks (15 + 15 + 15 +15) = 60
Section B: Answer any 2 of the 3 questions Each question counts 20 marks (2 x 20) = 40
Total = 100

SECTION A (60 marks)

Answer **ALL** four questions in section A

Section A requires brief and to the point answers

In most cases simply list, or briefly explain what is required

It may be advantageous to use statistical notation (mathematical symbols) to explain concepts, but make sure to also explain their meaning

It is not required to re-explain concepts that have been previously dealt with If required, you may simply refer to your previous answer/s

In general, each mark represents one correct fact or correct interpretation

You have 120 minutes to earn 100 marks in the case of the complete paper, that is, 6 minutes per 5 marks **SECTION A (60 marks)**

Answer **ALL** four questions in section A

QUESTION 1 (15 marks)

(a) List any three uses of econometrics (3)

(b) Briefly explain the meaning of the following

(i) Regression analysis

(ii) The estimated regression equation

(iii) Total, explained and residual sum of squares (7)

(c) Carefully distinguish between R^2 and $\overline{R}^2$ Can $\overline{R}^2$ be misused in regression analysis? Explain (5)

[15]

QUESTION 2 (15 marks)

(a) State four valid criteria for determining whether a given variable belongs in an equation (4)

(b) Distinguish between sequential specification searches and data mining (4)

(c) Econometricians are often careful to avoid specification bias (7)

What exactly is specification bias?

i What is known to cause specification bias?

ii Are unbiased estimates always better than biased estimates? Why or why not?

iii What's the best way to avoid specification bias?

[15]

QUESTION 3 (15 marks)

(a) What is the stochastic error term and why does it have to be included in a regression equation? (4)

(b) Briefly explain what OLS means and what it tries to achieve Why should researchers use it? (5)

(c) List the steps applied in regression analysis (6)

[15]

[TURN OVER]

QUESTION 4 (15 marks)

Suppose you are a junior researcher working for a research institute and your line manager gives you the following information that has been used to establish the impact of income on household consumption

$$\sum_{i=1}^{N} Y_i^2 = 23136,\ \sum_{i=1}^{N} X_i^2 = 76000,\ \sum_{i=1}^{N} X_i Y_i = 48925,\ \sum_{i=1}^{N} Y_i = 5650,\ \sum_{i=1}^{N} X_i = 2800,\ \sum_{i=1}^{N} x_i y_i = 1830$$

$$\sum_{i=1}^{N} y_i^2 = 3390,\ \sum_{i=1}^{N} x_i^2 = 2424,\ \sum_{i=1}^{N} e_i^2 = 3279,\ \bar{X} = 11\,2,\ \bar{Y} = 22\,6$$

Where $x_i = X_i - \bar{X}$ and $y_i = Y_i - \bar{Y}$ for $i = 1, 2, \quad , N$

(a) How many households were considered for this research analysis? (2)

(b) Use the equation $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ to compute the OLS estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ (3)

(c) What is the meaning of $\hat{\beta}_0$ and the slope coefficient $\hat{\beta}_1$ calculated in part (b)? (4)

(d) Compute the coefficient of determination (R^2) and the standard error of the slope coefficient SE($\hat{\beta}_1$) Do you think that the explanatory variable X is statistically significant? *Hint Use SE obtained to calculate the t-statistic* (6)

[15]

SUBTOTAL SECTION A **[60]**

SECTION B (40 marks)

Answer **ANY TWO** of the **THREE QUESTIONS** in section B Each question counts 20 marks

- Section B mostly consists of practical problems
- Use a 5% level of significance in statistical testing, unless stated otherwise

QUESTION B1 (20 marks)

A researcher has tried to establish the relationship between money supply (M3) and its two determinants, interest rates (IR) and economic growth (GDP) and estimated the following regression results

$$\hat{M}3_t = \underset{(0\ 2916)}{21\ 460} - \underset{(-0\ 107)}{0\ 4206}\, IR_t + \underset{(13\ 1632)}{0\ 7336}\, GDP_t$$

$R^2 = 0\ 9470$, $\bar{R}^2 = 0\ 9399$, Durbin-Watson (DW) statistic = 0 8049

Where M3 = money supply
IR = interest rate
GDP = gross domestic product

M and GDP are measured in millions and IR is measured in percentages The values in the parenthesis represent the t-statistics

(a) Fully interpret the above regression results, taking into account the units used Do you think the signs obtained conform to economic theory? (10)

(b) What econometric problem(s) does this equation appear to have? Explain (3)

(c) Suppose that you are now told that the simple correlation coefficient between IR and GDP is - 0 89 and the values of the variance inflation factors exceed 5 Does this change your answer to part (b) above? How? (3)

(d) What remedies for the problems you identify in parts (b) and (c) do you recommend? Explain (4)

QUESTION B2 (20 marks)

You are an economist working in the research department at your institution and in estimation the demand for beef, you obtained the following results The dependent variable is Q Per capita consumption of beef (kilograms per month)

Variable	Coefficient	Standard error	t-statistic
Constant	32 85	4 24	7 74
H Household income	0 0103	0 0053	194
PB price beef	-0 2979	0 1348	-2 21
PC price chicken	-0 0140	0 0642	-0 22
PS price of substitutes	0 1134	0 0820	1 38

Method Ordinary Least Squares, Number of observations in sample 23, R^2=0 92, Durbin-Watson statistic = 0 67, F = 51 8 The variables are as follows

H Per capita disposable household income (Rands)

PB Price of beef (cents/kg)

PC Price of chicken (cents/kg)

PS Price of beef substitutes (weighted average of chicken and pork, cents per kg)

(a) Interpret the meaning of the constant term and β_{PC} Use proper units of measurement if applicable (3)

(b) Explain how the t-value of -2 21 in the table has been calculated (1)

(c) Test each relevant coefficient for statistical significance State its expected value (on theoretical grounds), the null and alternative hypothesis and your conclusion (include the reason) Use a 5% level of significance for testing Please use a table to present your results (13)

(d) Identify a major shortcoming of this regression and explain what can be done to correct it (3)

QUESTION B3 (20 marks)

Regression results for the annual demand for money in country XYZ are given as follow

$$\hat{M}_t = 0\ 14 + 1\ 05\ Y_t^* + 0\ 01Y_t + 0\ 75IR_t$$
$$(0\ 16) \quad (0\ 11) \quad (0\ 07)$$
$$\bar{R}^2 = 0\ 986 \quad DW = 0\ 98 \quad N = 70 \quad r_{Y^*,Y} = 0\ 96$$

Where M_t = the log of the money stock
Y_t^* = the log of average income
Y_t = the log of current income
IR_t = the log of the rate of interest
$r_{y^*,y}$ = correlation coefficient between Y^* and Y
Values in parentheses are the standard errors

(a) Interpret the results given above (4)

(b) Hypothesise signs and test the appropriate null hypotheses at the 5% level of significance (9)

(c) What econometric problems seem likely to be in this equation? (3)

(d) In particular, are there are any problems related to the coefficient of Y? If so, are these problems more likely to have been caused by multicollinearity, serial correlation, or heteroskedasticity? (4)

Appendix 1: Formulae sheet

OLS estimates of $Y_t = \beta_0 + \beta_1 X_{1t} + \varepsilon_t$

$\hat{\beta}_1 = \dfrac{\sum x_{1t} y_t}{\sum x_{1t}^2}$ where

$SE(\hat{\beta}_1) = \sqrt{\dfrac{\left(\dfrac{\sum e_t^2}{n-2}\right)}{\sum x_{1t}^2}}$

$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1$ where $\overline{Y} = \sum_t Y_t / n$ and $\overline{X}_j = \sum_t X_{jt} / n$

TSS (Total sum of squares) = ESS (explained) + RSS (residual)

OLS estimates of $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$

$$\hat{\beta}_1 = \frac{\left(\sum y_t x_{1t}\right)\left(\sum x_{2t}^2\right) - \left(\sum y_t x_{2t}\right)\left(\sum x_{1t} x_{2t}\right)}{\left(\sum x_{1t}^2\right)\left(\sum x_{2t}^2\right) - \left(\sum x_{1t} x_{2t}\right)^2}$$

where

$x_{2t} = X_{2t} - \overline{X}_2$

$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1 - \hat{\beta}_2 \overline{X}_2$

$SE(\hat{\beta}_1) = \sqrt{\dfrac{\sum e_t^2 / (n-3)}{\left(\sum x_1^2\right)\left(1 - r_{X_1 X_2}^2\right)}}$

$$r_{X_1 X_2} = \frac{\sum \left(X_{1t} - \overline{X}_1\right)\left(X_{2t} - \overline{X}_2\right)}{\sqrt{\sum \left(X_{1t} - \overline{X}_1\right)^2 \left(X_{2t} - \overline{X}_2\right)^2}}$$

Some statistical measures

$t = \dfrac{\hat{\beta} - \beta_{H_o}}{SE(\hat{\beta})}$

$F = \dfrac{ESS / K}{RSS / (n - K - 1)}$

$DW\ d = \dfrac{\sum_{t=2}^{T} (e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$

Table	Contents
1	Critical values of the t-distribution
2	Critical values of the F-distribution 5% level of significance
3	Critical values of the Durbin-Watson test statistic (DW-d) of D_L and D_U DW-d 5% one-sided and 10% two-sided level of significance

Table 1: Critical values of the t-distribution

Degrees of freedom	Level of significance					
	One-sided	10%	5%	2 5%	1%	0 5%
	Two-sided	20%	10%	5%	2%	1%
1		3 078	6 314	12 706	31 821	63 657
2		1 886	2 920	4 303	6 965	9 925
3		1 638	2 353	3 182	4 541	5 841
4		1 533	2 132	2 776	3 747	4 604
5		1 476	2 015	2 571	3 365	4 032
6		1 440	1 943	2 447	3 143	3 707
7		1 415	1 895	2 365	2 998	3 499
8		1 397	1 860	2 306	2 896	3 355
9		1 383	1 833	2 262	2 821	3 250
10		1 372	1 812	2 228	2 764	3 169
11		1 363	1 796	2 201	2 718	3 106
12		1 356	1 782	2 179	2 681	3 055
13		1 350	1 771	2 160	2 650	3 012
14		1 345	1 761	2 145	2 624	2 977
15		1 341	1 753	2 131	2 602	2 947
16		1 337	1 746	2 120	2 583	2 921
17		1 333	1 740	2 110	2 567	2 898
18		1 330	1 734	2 101	2 552	2 878
19		1 328	1 729	2 093	2 539	2 861
20		1 325	1 725	2 086	2 528	2 845
21		1 323	1 721	2 080	2 518	2 831
22		1 321	1 717	2 074	2 508	2 819
23		1 319	1 714	2 069	2 500	2 807
24		1 318	1 711	2 064	2 492	2 797
25		1 316	1 708	2 060	2 485	2 787
26		1 315	1 706	2 056	2 479	2 779
27		1 314	1 703	2 052	2 473	2 771
28		1 313	1 701	2 048	2 467	2 763
29		1 311	1 699	2 045	2 462	2 756
30		1 310	1 697	2 042	2 457	2 750
40		1 303	1 684	2 021	2 423	2 704
50		1 299	1 676	2 009	2 403	2 678
60		1 296	1 671	2 000	2 390	2 660
70		1 294	1 667	1 994	2 381	2 648
120		1 289	1 658	1 980	2 358	2 617
Normal		1 282	1 645	1 960	2 326	2 576

Table 2: Critical values of the F-distribution: 5% level of significance

Degrees of freedom of denominator (v_2)	Degrees of freedom for numerator (v_1)											
	1	2	3	4	5	6	7	8	10	12	20	∞
1	161 0	200 0	216 0	225 0	230 0	234 0	237 0	239 0	242 0	244 0	248 0	254 3
2	18 5	19 0	19 2	19 2	19 3	19 3	19 4	19 4	19 4	19 4	19 4	19 5
3	10 1 3	9 55	9 28	9 12	9 01	8 94	8 89	8 85	8 79	8 74	8 66	8 53
4	7 71	6 94	6 59	6 39	6 26	6 16	6 09	6 04	5 96	5 91	5 80	5 63
5	6 61	5 79	5 41	5 19	5 05	4 95	4 88	4 82	4 74	4 68	4 56	4 36
6	5 99	5 14	4 76	4 53	4 39	4 28	4 21	4 15	4 06	4 00	3 87	3 67
7	5 59	4 74	4 35	4 12	3 97	3 87	3 79	3 73	3 64	3 57	3 44	3 23
8	5 32	4 46	4 07	3 84	3 69	3 58	3 50	3 44	3 35	3 28	3 15	2 93
9	5 12	4 26	3 86	3 63	3 48	3 37	3 29	3 23	3 14	3 07	2 94	2 71
10	4 96	4 10	3 71	3 48	3 33	3 22	3 14	3 07	2 98	2 91	2 77	2 54
11	4 84	3 98	3 59	3 36	3 20	3 09	3 01	2 95	2 85	2 79	2 65	2 40
12	4 75	3 89	3 49	3 26	3 11	3 00	2 91	2 85	2 75	2 69	2 54	2 30
13	4 67	3 81	3 41	3 18	3 03	2 92	2 83	2 77	2 67	2 60	2 46	2 21
14	4 60	3 74	3 34	3 11	2 96	2 85	2 76	2 70	2 60	2 53	2 39	2 13
15	4 54	3 68	3 29	3 06	2 90	2 79	2 71	2 64	2 54	2 48	2 33	2 07
16	4 49	3 63	3 24	3.01	2 85	2 74	2 66	2 59	2 49	2 42	2 28	2 01
17	4 45	3 59	3 20	2 96	2 81	2 70	2 61	2 55	2 45	2 38	2 23	1 96
18	4 41	3 55	3 16	2 93	2 77	2 66	2 58	2 51	2 41	2 34	2 19	1 92
19	4 38	3 52	3 13	2 90	2 74	2 63	2 54	2 48	2 38	2 31	2 16	1 88
20	4 35	3 49	3 10	2 87	2 71	2 60	2 51	2 45	2 35	2 28	2 12	1 84
21	4 32	3 47	3 07	2 84	2 68	2 57	2 49	2 42	2 32	2 25	2 10	1 81
22	4 30	3 44	3 05	2 82	2 66	2 55	2 46	2 40	2 30	2 23	2 07	1 78
23	4 28	3 42	3 03	2 80	2 64	2 53	2 44	2 37	2 27	2 20	2 05	1 76
24	4 26	3 40	3 01	2 78	2 62	2 51	2 42	2 36	2 25	2 18	2 03	1 73
25	4 24	3 39	2 99	2 76	2 60	2 49	2 40	2 34	2 24	2 16	2 01	1 71
26	4 23	3 37	2 98	2 74	2 59	2 47	2 39	2 32	2 22	2 15	1 99	1 69
27	4 21	3 35	2 96	2 73	2 57	2 46	2 37	2 31	2 20	2 13	1 97	1 68
28	4 20	3 34	2 95	2 71	2 56	2 45	2 36	2 29	2 19	2 12	1 96	1 66
29	4 18	3 33	2 93	2 70	2 55	2 43	2 35	2 28	2 18	2 10	1 94	1 64
30	4 17	3 32	2 92	2 69	2 53	2 42	2 33	2 27	2 16	2 09	1 93	1 62
40	4 08	3 23	2 84	2 61	2 45	2 34	2 25	2 18	2 08	2 00	1 84	1 51
50	4 03	3 18	2 79	2 56	2 40	2 29	2 20	2 13	2 03	1 95	1 78	1 44
60	4 00	3 15	2 76	2 53	2 37	2 25	2 17	2 10	1 99	1 92	1 75	1 39
70	3 98	3 13	2 74	2 50	2 35	2 23	2 14	2 07	1 97	1 89	1 72	1 36
120	3 92	3 07	2 68	2 45	2 29	2 18	2 09	2 02	1 91	1 83	1 66	1 25
∞	3 84	3 00	2 60	2 37	2 21	2 10	2 01	1 94	1 83	1 75	1 57	1 00

Table 3: Critical values of the Durbin-Watson test statistics D_L and D_U 5% one-sided and 10% two-sided level of significance

n	k'=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1 08	1 36	0 95	1 54	0 81	1 75	0 69	1 97	0 56	2 21	0 45	2 47	0 34	2 73
16	1 11	1 37	0 98	1 54	0 86	1 73	0 73	1 93	0 62	2 15	0 50	2 39	0 40	2 62
17	1 13	1 38	1 02	1 54	0 90	1 71	0 78	1 90	0 66	2 10	0 55	2 32	0 45	2 54
18	1 16	1 39	1 05	1 53	0 93	1 69	0 82	1 87	0 71	2 06	0 60	2 26	0 50	2 46
19	1 18	1 40	1 07	1 53	0 97	1 68	0 86	1 85	0 75	2 02	0 65	2 21	0 55	2 40
20	1 20	1 41	1 10	1 54	1 00	1 68	0 89	1 83	0 79	1 99	0 69	2 16	0 60	2 34
21	1 22	1 42	1 13	1 54	1 03	1 67	0 98	1 81	0 83	1 96	0 73	2 12	0 64	2 29
22	1 24	1 43	1 15	1 54	1 05	1 66	0 96	1 80	0 86	1 94	0 77	2 09	0 68	2 25
23	1 26	1 44	1 17	1 54	1 08	1 66	0 99	1 79	0 90	1 92	0 80	2 06	0 72	2 21
24	1 27	1 45	1 19	1 55	1 10	1 66	1 01	1 78	0 93	1 90	0 84	2 04	0 75	2 17
25	1 29	1 45	1 21	1 55	1 12	1 66	1 04	1 77	0 95	1 89	0 87	2 01	0 78	2 14
26	1 30	1 46	1 22	1 55	1 14	1 65	1 06	1 76	0 98	1 88	0 90	1 99	0 82	2 12
27	1 32	1 47	1 24	1 56	1 16	1 65	1 08	1 76	1 00	1 86	0 93	1 97	0 85	2 09
28	1 33	1 48	1 26	1 56	1 18	1 65	1 10	1 75	1 03	1 85	0 95	1 96	0 87	2 07
29	1 34	1 48	1 27	1 56	1 20	1 65	1 12	1 74	1 05	1 84	0 98	1 94	0 90	2 05
30	1 35	1 49	1 28	1 57	1 21	1 65	1 14	1 74	1 07	1 83	1 00	1 93	0 93	2 03
31	1 36	1 50	1 30	1 57	1 23	1 65	1 16	1 74	1 09	1 83	1 02	1 92	0 95	2 02
32	1 37	1 50	1 31	1 57	1 24	1 65	1 18	1 73	1 11	1 82	1 04	1 91	0 97	2 00
33	1 38	1 51	1 32	1 58	1 26	1 65	1 19	1 73	1 13	1 81	1 06	1 90	0 99	1 99
34	1 39	1 51	1 33	1 58	1 27	1 65	1 21	1 73	1 14	1 81	1 08	1 89	1 02	1 98
35	1 40	1 52	1 34	1 58	1 28	1 65	1 22	1 73	1 16	1 80	1 10	1 88	1 03	1 97
36	1 41	1 52	1 35	1 59	1 30	1 65	1 24	1 73	1 18	1 80	1 11	1 88	1 05	1 96
37	1 42	1 53	1 36	1 59	1 31	1 66	1 25	1 72	1 19	1 80	1 13	1 87	1 07	1 95
38	1 43	1 54	1 37	1 59	1 32	1 66	1 26	1 72	1 20	1 79	1 15	1 86	1 09	1 94
39	1 43	1 54	1 38	1 60	1 33	1 66	1 27	1 72	1 22	1 79	1 16	1 86	1 10	1 93
40	1 44	1 54	1 39	1 60	1 34	1 66	1 29	1 72	1 23	1 79	1 18	1 85	1 12	1 93
45	1 48	1 57	1 43	1 62	1 38	1 67	1 34	1 72	1 29	1 78	1 24	1 84	1 19	1 90
50	1 50	1 59	1 46	1 63	1 42	1 67	1 38	1 72	1 34	1 77	1 29	1 82	1 25	1 88
55	1 53	1 60	1 49	1 64	1 45	1 68	1 41	1 72	1 37	1 77	1 33	1 81	1 29	1 86
60	1 55	1 62	1 51	1 65	1 48	1 69	1 44	1 73	1 41	1 77	1 37	1 81	1 34	1 85
65	1 57	1 63	1 54	1 66	1 50	1 70	1 47	1 73	1 44	1 77	1 40	1 81	1 37	1 84
70	1 58	1 64	1 55	1 67	1 53	1 70	1 49	1 74	1 46	1 77	1 43	1 80	1 40	1 84
75	1 60	1 65	1 57	1 68	1 54	1 71	1 52	1 74	1 49	1 77	1 46	1 80	1 43	1 83
80	1 61	1 66	1 59	1 69	1 56	1 72	1 53	1 74	1 51	1 77	1 48	1 80	1 45	1 83
85	1 62	1 67	1 60	1 70	1 58	1 72	1 55	1 75	1 53	1 77	1 50	1 80	1 47	1 83
90	1 63	1 68	1 61	1 70	1 59	1 73	1 57	1 75	1 54	1 78	1 52	1 80	1 49	1 83
95	1 64	1 69	1 62	1 71	1 60	1 73	1 58	1 75	1 56	1 78	1 54	1 80	1 51	1 83
100	1 65	1 69	1 63	1 72	1 61	1 74	1 59	1 76	1 57	1 78	1 55	1 80	1 53	1 83

n=number of observations
k'=number of explanatory variables excluding the constant term

ROUGH WORK WILL NOT BE MARKED

ROUGH WORK WILL NOT BE MARKED

ROUGH WORK WILL NOT BE MARKED

ROUGH WORK WILL NOT BE MARKED

ROUGH WORK WILL NOT BE MARKED