UNIVERSITY OF ESWATINI

SECOND SEMESTER MAIN EXAMINATION PAPER, NOVEMBER 2021

FACULTY OF SOCIAL SCIENCES

DEPARTMENT OF ECONOMICS

COURSE CODE: ECO420

TITLE OF PAPER: ECONOMETRIC METHODS II

TIME ALLOWED: 2 HOURS

Instructions

1. This paper consists of two (2) sections, A and B
2. Section A, is compulsory and carries 40 marks
3. Section B, contains three (3) questions
4. Answer any other two (2) questions in Section B

Special Requirements

Scientific calculator

Additional Material (s)

None

DO NOT turn examination paper over until instructed to do so.

SECTION A

Question 1 Compulsory [40]

a) Table 1 presents data on management fees that a leading mutual fund firm in the US pays to its investment advisors to manage its assets. The fees depend on the net asset value of the fund. The higher the net asset value of the fund the lower the advisory fees. The distribution of management fees can be assumed to follow an exponential growth function $Y_i = \beta_1 e^{\beta_2 X_i + u_i}$. Use direct optimization to determine the estimates that minimize the sum of squares of the error. [15]

Table 1

	Fee %	Net Asset Value in Billions
1	0.52	0.5
2	0.508	5
3	0.484	10
4	0.46	15
5	0.4398	20
6	0.4238	25
7	0.4115	30
8	0.402	35
9	0.3944	40
10	0.388	45
11	0.3825	55
12	0.3738	60

b) When J=2, we form J – 1 = 1 non redundant logits, when J > 2, we can form J – 1 non redundant logits. Prove this statement. [7]

c) Suppose we have count data which can be modelled using the Poisson Distribution with parameters (λ). Prove that $\lambda = E[X] = E[X - E[X]]^2$ [8]

d) Suppose we want to estimate the association of Aspirin Use on Myocardial Infarction. Calculate the generalized odds ratio, random row or column odds ratio and the reference category odds ratio and interpret the result. [6]

Table 2

	Fatal Attack	Myocardial Infarction Nonfatal Attack	No Attack
Placebo	18	171	10845
Aspirin	5	99	10933

e) What is panel data? [2]

f) What are the drawbacks of relying on the coefficient of determination? [2]

SECTION B

ANSWER ANY TWO QUESTIONS

[30]

Question 2

a) Suppose you are a research consultant hired by the Deputy Prime Minister's Office to conduct a study on employment sector preferences by sex. The aim is to find out which characteristics determine employment in the public sector, private sector and the informal sector. The following output is obtained from running the data.

Table 2: Model Fitting Information

	Women	Men
Number of observations	1567	2043
LR chi2(38)	1077.42	828.1
Prob > chi^2	0.0000	0.0000
Pseudo R^2	0.3342	0.241

Table 3: Test for Independence of Irrelevant Alternatives (Females)

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
Formal Public Sector	-301.809	-296.145	11.328	20	0.937	for Ho
Informal Private Sector	-217.252	-207.92	18.664	20	0.544	for Ho
Formal Private Sector	-109.108	-97.256	23.704	20	0.256	for Ho

Table 4: Test for Independence of Irrelevant Alternatives (Males)

Omitted	lnL(full)	lnL(omit)	chi2	df	P>chi2	evidence
Formal Public Sector	-276.608	-271.078	11.059	20	0.945	for Ho
Informal Private Sector	-339.212	-329.412	19.601	20	0.483	for Ho
Formal Private Sector	-134.405	-125.092	18.626	20	0.546	for Ho

Table 5: Multinomial Logit Model Results (Female)

	Formal Public Sector		Informal Private Sector	
	Coefficient	**Odds Ratio**	**Coefficient**	**Odds Ratio**
age	0.0521***	1.0534	0.0463***	1.0474
Administrative Region				
Manzini	-0.0436	0.9573	0.3787**	1.4604
Shiselweni	0.6267**	1.8714	0.2727	1.3135
Lubombo	-0.3187	0.7271	0.2042	1.2265
Geographic Location				
Urban	-0.9692***	0.3794	-0.8385***	0.4324
Marital Status				

Married	0.8247***	2.2812	0.5103***	1.6658
Widowed	-0.5322	0.5873	0.4636*	1.5898
Divorced	0.3264	1.3860	0.4403	1.5532
Household Position				
Head	0.4939**	1.6388	0.1172	1.1244
Educational Attainment				
Elementary	0.7453	2.1071	0.1633	1.1774
High school	2.1502***	8.5864	-0.7444***	0.4750
Vocational and College	3.4309***	30.9037	-1.7894***	0.1671
University	3.3563***	28.6834	-34.3019	1.27E-15
Industrial Classification				
Agriculture and mining	-2.7905***	0.0614	-0.5437	0.5806
Manufacturing	-2.8310***	0.0590	-0.2932	0.7459
Electricity and water	-0.6854	0.5039	-33.1531	4.00E-15
Construction	0.6803	1.9744	1.2305	3.4230
Wholesale, retail trade and hospitality industry	-3.5288***	0.0293	1.3061***	3.6920
Transportation and storage	-0.2749	0.7596	0.2183	1.2440
constant	-4.2101***		-2.3160***	

Significant at 10%, **Significant at 5%, * Significant at 1%*

Table 6: Multinomial Logit Results Male

	Formal Public Sector		Informal Private Sector	
	Coefficient	**Odds Ratio**	**Coefficient**	**Odds Ratio**
age	0.0215***	1.0217	0.0157**	1.0158
Administrative Region				
Manzini	0.0356	1.0363	0.2223	1.2489
Shiselweni	0.7093***	2.0325	0.2856	1.3306
Lubombo	-0.3155	0.7294	-0.3059	0.7365
Geographic Location				
Urban	-0.7578***	0.4687	-0.6159***	0.5402
Marital Status				
Married	0.3748**	1.4548	-0.4305**	0.6502
Widowed	0.1237	1.1317	-0.4054	0.6667
Divorced	0.2629	1.3007	-0.1384	0.8708
Household Position				
Head	0.0576	1.0593	0.1784	1.1953
Educational Attainment				
Elementary	-0.1973	0.8210	-0.3380	0.7132
High School	0.7143***	2.0427	-0.4510*	0.6370

Vocational and College Education	1.6860***	5.3981	-0.8240*	0.4387
University	1.7338***	5.6619	-1.8534**	0.1567
Industrial Classification				
Agriculture and mining	-1.9861***	0.1372	-1.6642***	0.1893
Manufacturing	-3.1148***	0.0444	-1.3683***	0.2545
Electricity and water	0.4582	1.5812	0.3352	1.3983
Construction	-2.1335***	0.1184	0.4884**	1.6298
Wholesale, retail trade and hospitality industry	-3.4851***	0.0307	0.1859	1.2043
Transportation and storage	-1.7685***	0.1706	-0.8747***	0.4170
constant	-1.4302***		-1.1551***	

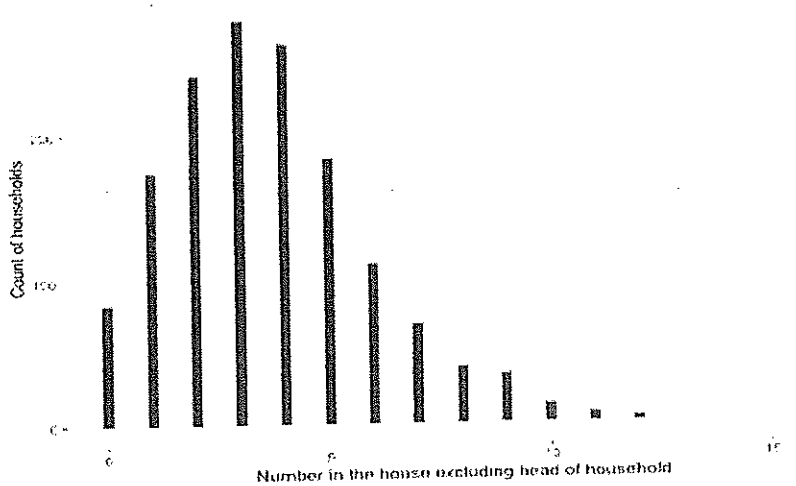
Significant at 10%, **Significant at 5%, * Significant at 1%*

i. Based on the above information, what type of model was fitted? [2]
ii. What was the reference category? [3]
iii. Comment on the model fitting information and what it means about the sufficiency of the model. [5]
iv. What are the assumptions of this model? Do they hold for the particular dataset? [5]
v. Interpret the results of the study for both males and females. [15]

[30]

Question 3

a) Suppose that the government is interested in determining the age at which heads of households in Eswatini most likely to find the largest number of people in their household. Specify the appropriate Poisson regression model for this problem. [3]
b) The distribution of number of people per household is plotted below, comment on the likely distribution of the number of people per household. [4]



c) Given the following array of means and variances what does the data imply in terms of the assumptions of the Poisson regression model? [6]

Age Groups	Mean	Variance	n
(15,20]	1.666667	0.6666667	6
(20,25]	2.166667	1.5588235	18
(25,30]	2.918367	1.4098639	49
(30,35]	3.444444	2.1931464	108
(35,40]	3.841772	3.5735306	158
(40,45]	4.234286	4.4447947	175
(45,50]	4.489691	6.3962662	194
(50,55]	4.010638	5.2512231	188
(55,60]	3.806897	6.5318966	145
(60,65]	3.705882	6.1958204	153
(65,70]	3.339130	7.9980168	115
NA	2.549738	5.5435657	191

d) The following model was run using R, interpret the results. [7]

$$\widehat{log(\lambda)} = 1.55 - 0.0047(age)$$

```
glm(formula = total ~ age, family = poisson, data = fHH1)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.5499422  0.0502754  30.829  < 2e-16 ***
age         -0.0047059  0.0009363  -5.026 5.01e-07 ***
---
(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2362.5  on 1499  degrees of freedom
Residual deviance: 2337.1  on 1498  degrees of freedom
AIC: 6714
```

e) The above model in (d) points into a situation of likely over dispersion in the model. What are the likely causes for over dispersion and what are the likely remedial measures? [5]

f) Highlight the method of recursive least squares for testing for structural breaks in a dataset. [5]

[30]

Question 4

Figure 1 presents regressions that were estimated using data for 48 US states from 1982 – 1988 (336 observations in total). The study aimed at ascertaining the impact of socio-economic conditions and

traffic laws on the number of traffic related death accidents. Standard errors are given in parentheses under the coefficients and the p-values are given in parenthesis under the F-statistic. Maximum allowable type 1 error is 5%.

a) Which is the base model? Why? [2]
b) Compare and contrast the six models. [6]
c) Interpret the various F statistics for the models, also comment on which model will be adequate for estimating the impact of laws and socio-economic status on traffic accidents. [8]
d) Interpret the results of the model and clearly state which factors influence traffic related deaths in the US states. [10]

Dependent Variable: Traffic Fatality Rate (Deaths Per 10,000).

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Beer tax	0.36** (0.05)	−0.66** (0.20)	−0.64* (0.25)	−0.45* (0.22)	−0.70** (0.25)	−0.46* (0.22)
Drinking age 18				0.028 (0.036)	−0.011 (0.064)	
Drinking age 19				−0.019 (0.040)	−0.078 (0.049)	
Drinking age 20				0.031 (0.046)	−0.102* (0.046)	
Drinking age						−0.002 (0.017)
Mandatory jail?				0.013 (0.032)	−0.026 (0.065)	
Mandatory community service?				0.033 (0.115)	0.197 (0.137)	
Mandatory jail or community service?						0.031 (0.076)
Average vehicle miles per driver				0.008 (0.008)	0.017 (0.010)	0.009 (0.008)
Unemployment rate				−0.063** (0.012)		−0.063** (0.012)
Real income per capita (logarithm)				1.81** (0.47)		1.79** (0.45)
State effects?	no	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes
F-statistics and p-values Testing Exclusion of Groups of Variables:						
Time effects = 0			2.47 (0.024)	11.44 (<0.001)	2.28 (0.037)	11.50 (<0.001)
Drinking age coefficients = 0				0.48 (0.696)	2.09 (0.102)	
Jail, community service coefficients = 0				0.17 (0.845)	0.59 (0.557)	
Unemployment rate, income per capita = 0				38.29 (<0.001)		40.12 (<0.001)
$\bar{R}^2$	0.090	0.889	0.891	0.926	0.893	0.926

Figure 1: Regression Analysis of the Effect of Drunk Laws on Traffic Deaths

e) Distinguish between intrinsically linear models and non-linear models. [4]