Task in econometrics I

7. Models with qualitative variables

7.1 For 6 consumers the demand for protein shakes (y), income (x) and age group (a) were surveyed:

Consumers	1	2	3	4	5	6
Demand for protein shakes	8	12	6	15	11	9
income	10	18	8	22	14	12
ager	jung	alt	jung	alt	jung	jung

a) Estimate the demand function for protein shakes considering the age effect,

$$y_t = e^{\beta_1} \cdot x_t^{\beta_2} \cdot e^{\beta_3 \cdot a_t} \cdot e^{u_t}$$

with

$$a_t = \begin{cases} 0, & \text{falls Konsument jung} \\ 1, & \text{falls Konsument alt} \end{cases},$$

in linearized form with the ordinary least squares method (OLS) using the program R! Give the linearized demand function used for this purpose and the corresponding regression output!

b) Specify the linearized OLS estimated demand function for protein shakes and interpret the income effect!!

c) What are the linearized age-specific demand functions for protein shakes? Interpret the age effect on the demand for protein shakes!

7.2 Data are given for 6 consumers on the demand for protein shakes (y), income (x) and age group (a) (see 7.1).

a) How must the multiplicative demand function for protein shakes given in 7.1a) be modified to reflect the age-dependent influence on income elasticity?

b) Specify the linearized demand function for protein shakes, taking into account the interaction between income and age, and estimate it using the ordinary least squares method (OLS method)! Give the determined regression output of the program R!

c) What is the numerically specified demand function for protein shakes! Interpret the income and age effect on the demand for protein shakes by specifying the age-specific demand functions!

7.3 The production y of an industrial sector in a quarter can be traced back to the order intake x delayed by one period and to additive seasonal influences. Give the corresponding design of the observation matrix X for a period of three years!

7.4 For the Keynesian consumption function, EViews produced the following estimated results for the total period 1974 -1992,

Dependent Variable: KONSUMR
Method: Least Squares Sample: 1974 1992

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	37.31676	44.21441	0.843995	0.4104
VEINKR	0.864223	0.037617	22.97439	0.0000
R-squared	0.968797	S.D. dependent var		128.5075
Adjusted R-squared	0.966962	S.E. of regression		23.35807

before

Dependent Variable: KONSUMR
Method: Least Squares Sample: 1974 1979

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-118.7550	56.85496	-2.088736	0.1050
VEINKR	0.997260	0.054577	18.27246	0.0001
R-squared	0.988162	S.D. dependent var		63.73863
Adjusted R-squared	0.985202	S.E. of regression		7.753625

and after the oil price shock 1980:

Dependent Variable: KONSUMR
Method: Least Squares Sample: 1980 1992

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	167.1905	47.73910	3.502170	0.0050
VEINKR	0.764885	0.038743	19.74249	0.0000
R-squared	0.972553	S.D. dependent var		106.1633
Adjusted R-squared	0.970057	S.E. of regression		18.37044

Use the Chow test to check whether the 1980 oil price shock could have caused a structural break in the Keynesian consumption function ($\alpha = 0.01$)!

7.5 Interpret the EViews result of an OLS estimate of the Keynesian consumption function using the dummy variables D_1 (1 for 1974 to 1979), 0 else and D_2 (1 for 1980 to 1992, 0 else):

Dependent Variable: KONSUMR
Method: Least Squares Sample: 1974 1992

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D1	-118.7550	119.0320	-0.997673	0.3343
D2	167.1905	42.18468	3.963298	0.0012
VEINKR*D1	0.997260	0.114263	8.727736	0.0000
VEINKR*D2	0.764885	0.034235	22.34197	0.0000

7.6 What objections can be raised to the linear probability model for qualitative dependent variables from an econometric perspective?

7.7 X_i, i=1,2,…,n is a Bernoulli-distributed random variable,which takes the value 1 with probability p and the value 0 with probability 1-p. For the given random sample (x_1, x_2, …, x_n) from a Bernoulli distributed population,determine the maximum likelihood estimator (ML-estimator) for the probability p!

7.8 The choice of a means of transport V (public transport or private car) for the way to work has been estimated with EViews on the basis of the Logit model as a function of the distance E (in km) and the occupational status S using the maximum likelihood method (data see appendix):

Dependent Variable: VERKEHRSART
Method: ML - Binary Logit (Newton-Raphson)
Sample: 1 24 Included observations: 24
Convergence achieved after 6 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-5.005772	2.122852	-2.358041	0.0184
ENTFERNUNG	0.265619	0.134538	1.974306	0.0483
STELLUNG	4.590040	1.957320	2.345063	0.0190

R-squared	?	Mean dependent var	0.416667
S.D. dependent var	0.503610	S.E. of regression	0.280037
Akaike info criterion	0.744953	Sum squared resid	1.646835
Schwarz criterion	0.892210	Log likelihood	-5.939442
Restr. log likelihood	-16.30064	LR statistic	20.72239
Avg. log likelihood	-0.247477	Prob(LR statistic)	3.16E-05
Obs with Dep=0	14	Total obs	24
Obs with Dep=1	10		

a) Enter the numerically specified logit model of the means of transport selection

- with the estimated probabilities $\hat{p}_t$,

- with the estimated log odds $\ln[\hat{p}_t/(1-\hat{p}_t)]$

as regression values!

b) How can the effect coefficients of the variables distance (x_2) and occupational status (x_3) be interpreted?

c) Specify the marginal effect of distance to work (x_2) on the probability of choosing a car for employees who have the profile

- of the employee1,

- of the employee 2

own?

d) What is the size of the pseudo-R^2 marked „R-squared" in EViews? Like can it be interpreted?

Appendix: Data on choice of transport mode (Task 7.6)

Employee (t)	Type of transport (y)	distance (x_2)	Professional position(x_3)
1	1	26	1
2	0	5	0
3	1	17	1
4	0	1	1
5	0	14	0
6	1	28	1
7	0	19	0
8	0	2	0
9	1	22	0
10	0	8	0
11	0	6	0
12	1	18	1
13	0	4	0
14	0	12	0
15	0	3	0
16	1	15	1
17	1	8	0
18	0	10	0
19	0	2	0
20	1	10	1
21	1	16	1
22	0	11	0
23	0	4	0
24	1	5	1

Type of transport (y): 0 PT, 1 PC

Professional position (x_2): 0 low professional position, 1 high professional position