42178 – Transport System Analysis E21
Portfolio exercise I

The page limit of the full exercise is 15 pages, while the page limit on Exercise 3 is four pages excluding references. Feel free to put additional material in appendices but please keep essential results within the text so that your answers are understandable without appendices.

Exercise 1

The data ”cities_data.xlsx” include the annual driving of 4562 households and information about these households and the cities in which they live. The data include the following variables:

Variable	Variable description
VKMS	Vehicle km (annual)
LN_VKMS	LN(VKMS)
Car	Number of cars
Elderly	Dummy for elderly person in household
Worker	Dummy for working adult in household
Children	Number of children
HHincome	Household income in 1000 DKK
LN_INC	LN(HHincome)
RoadDens	Road Density in km/km2
TransSup	Supply of mass transit km/km2
CityShape	City shape (1=circular city, ~0 very narrow city)

Table 1: Variables in the cities data set.

a) Make descriptive statistics, e.g. a table with summary statistics (min, mean, std. dev., and max) and a correlation matrix, for the variables that you are going to use and discuss the results (read the whole exercise first so that you know which variables you have to use).
b) Estimate a linear regression explaining the number of vehicle km either in total or per vehicle using relevant socio-economic variables as explanatory variables. Discuss the results.
c) Add the supply and city variables to the model and redo the estimation. Discuss the results.
d) Compare the model from b) with the model from c). What model do you prefer and why?
e) Calculate the average elasticity of VKMS with respect to household income and mass transit supply. Interpret the elasticity that you find.
f) Calculate the effect on the driving across the cities in a future scenario where all cities are expected to increase their mass transit supply with 50% while income is expected to increase by 10% (all remaining variables are assumed to be unchanged).
g) Assuming your model to be correct, briefly discuss uncertainties in your forecast in f).

Exercise 2

The data ”freqdata.xlsx” include the number of casualties in traffic among 16 years old in 20 American states across 20 years. The data include the following variables:

Variable	Variable description
Fatal	number of fatalities among 16 years old
Year	year of the observation
State	number of the US state where the fatalities occurred
lnpop	logarithm of the population of the state in the period
unempl	unemployment rate of the state in the period
gdl	1 = active graduated driver licensing law, 0 = otherwise
Beltsc	1 = active secondary seat belt law, 0 = otherwise
Beltpr	1 = active primary seat belt law, 0 = otherwise
Bac08	1 = active blood alcohol concentration limit equal to 0.08 law, 0 = otherwise
zerotol	1 = active zero tolerance on alcohol law, 0 = otherwise
Alr	1 = active license revocation law, 0 = otherwise
Sp70	1 = active 70mph law on motorways, 0 = otherwise
Sp65	1 = active 65mph law on motorways, 0 = otherwise

Table 2: Variables in the freqdata data set.

a) Make descriptive statistics, e.g. a table with summary statistics and a correlation matrix, and discuss the results, for the variables that you are going to use (read the whole exercise first to decide which variables are relevant in the descriptive statistics).
b) Table 3 shows the estimation results from a Poisson regression with the number of accidents as dependent variable and using lnPop, unempl and sp65 as explanatory variables. Comment on the results.

Parameter	Estimate	Std. error	z/t statistic	P value
Intercept	-5.47	0.14	-38.9	<0.001
lnPop	0.75	0.01	57.4	<0.001
Unemployment	-0.09	0.01	-11.9	<0.001
Sp65	-0.44	0.02	-19.7	<0.001
No. obs.	960			
No. par.	4			
Null LL(intercept only)	-5026.9			
Final LL	-3124.2			
Adj. rho sq.	0.378			

Table 3: Model 1 estimation results based on the freqdata data set.

c) Table 4 shows the estimation results from two other Poisson regressions with the number of accidents as dependent variable and additional explanatory variables. Comment on the

differences among the models in Tables 3 and 4 and argue which model you prefer, e.g. by the use of LR tests as well as by looking at the signs of parameters.

Parameter	Estimate	z/t statistic	Estimate	z/t statistic
Intercept	-5.66	-36.4	-5.62	-36.4
lnPop	0.77	54.9	0.78	56.5
Unemployment	-0.08	-10.0	-0.09	-11.0
Sp65	-0.37	-12.9	-0.35	-13.8
gdl	-0.05	-1.5		
Bac08	-0.21	-6.6	-0.09	-3.56
Beltsc	-0.18	-5.0	-0.21	-6.0
Beltpr	-0.20	-7.5	-0.21	-8.1
Zerotol	0.05	1.7		
Alr	0.22	8.2		
No. obs.	960		960	
No. par.	10		7	
Null LL(intercept only)	-5026.9		-5026.9	
Final LL	-3030.1		-3068.5	
Adj. rho sq.	0.395		0.388	

Table 4: Model 2 and 3 estimation results based on the freqdata data set.

d) Based on your preferred model, calculate the min, max and average sample elasticity of expected fatalities with respect to the two variables population and unemployment rate. Comment your results.
e) Find the effect on the expected number of fatalities of Sp65, Bac08 and Beltpr based on your final model. Discuss the results.
f) Assuming your model to be correct, briefly discuss uncertainties in your elasticities and effect calculation in d) and e).

Exercise 3 (not part of oral exam but assessed on a scale 0-4 that is included in the final overall assessment)

In this third exercise, we ask you to analyse the transport system within a city of your choice. The aim is to make you think about issues that are relevant for the analysis of transport systems. To do the exercise you should

1) Select a city of your own choice in which you would like to analyse the transport system
2) Identify the most important aspects of the system
3) Identify one major sustainability issue in the system and discuss how the system could be improved to diminish this problem

You can get inspiration in

1) JR chapter 1
2) Hanna et al. (2017) Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta
3) Axsen et al. (2020) Crafting strong, integrated policy mixes for deep CO_2 mitigation in road transport.
4) Other literature that you find relevant
5) The real world

The page limit for Exercise 3 is four pages excluding reference. Focus on making your arguments as clear and concise as possible and keep the written text well structured.

Feel free to discuss your choice of transport system with Stefan or Mirka to get comments.