ECON4137 - Spring 2022

IMPORTANT: Always motivate your answers. Show knowledge and understanding of the concepts taught in the course. Your answers should be as short as possible and as long as necessary. Total marks of the examination are 170. Each subquestion carries 10 marks.

---

1. Suppose you are handed a dataset with n independently identically distributed variables, Y_i. The variables only take on positive values. Assume that the draws come from a distribution with the following density:

$$f(y) = \frac{\alpha y^{\alpha-1}}{\theta^{\alpha}} \exp\left(-\left(\frac{y}{\theta}\right)^{\alpha}\right)$$

Suppose that α is known and larger than zero, but θ is unknown and therefore needs to be estimated.

(a) In words describe the logic behind maximum likelihood estimation. Under what condition is the maximum likelihood estimator consistent and efficient? Is is generally unbiased?

(b) Derive the log-likelihood function

(c) Solve for the estimator, $\hat{\theta}$. Does this estimator maximize the likelihood?

(d) You are given the following sample of Y_i in Table 1. Assume that $\alpha = 2$. What is the maximum-likelihood estimate of θ? Explain, briefly, how you would calculate the variance of the estimate. You are *not* required to calculate the estimate of the variance.

Table 1: Sample of Y_i

	1	2	3	4	5	6	7	8	9	10
Y_i	6.2	4.0	5.2	9.6	0.5	2.5	7.4	9.3	2.7	6.5

2. A government agency wants to know the effect of providing training to unemployed. They therefore run an experiment, where they allocate a group of individuals randomly into two groups. A treatment group ($G_i = 1$) receives an offer to enter a training program while a control group ($G_i = 0$) does not. The agency wishes to measure the effect of the training program on an outcome Y_i. The government also observes whether the applicants participated in the training program ($T_i = 1$) or not ($T_i = 0$).

   (a) What kind of experiment is this? What assumptions are required for estimating a causal effect of the being offered the program?

   (b) The agency asks you to analyze the experiment. Using the variance-covariance matrix in Table 2 calculate the effect of offering the program to the unemployed. (Note: You may ignore the unit of measurement of the outcome)

   (c) Some of the individuals in the control group ended up participating in the training program, and some in the treatment group chose not to participate in the training program. In this light how do you interpret the effect that you estimated in question 2.b?

   (d) What share of the population comply with the assignment to treatment and control group, i.e. take up treatment if it's offered to them?

   (e) What is the effect of the training program for those who got it because the were allocated to the treatment group? What is this effect usually called? How does it differ from the average treatment effect of the treated and do they ever coincide?

Table 2: Variance-Covariance matrix

	Y_i	T_i	G_i
Y_i	1.04	0.24	0.07
T_i	0.24	0.21	0.06
G_i	0.07	0.06	0.25

3. During the Covid-19 pandemic policymakers have implemented a wide range of non-pharmaceutical interventions to reduce the number of infections. An important question is whether such policies effectively reduce infections.

   (a) Many countries decided to implement a nationwide lockdown in March 2020. However, Swedish officials chose not to implement a nationwide lockdowns. You have access to weekly infection rates for multiple countries. Your colleague suggest comparing Sweden to countries without lockdowns to estimate the effect of the lockdowns by running the following regression on data from the first half of 2020:

   $$Y_{it} = \beta_0 + \beta_1 Sweden_i + u_{it},$$

   where Y_{it} denotes infection cases in country i at week t and $Sweden_i$ is a dummy which is 1 if country i is Sweden. What would β_1 capture? Does this capture the effect of lockdowns? Discuss why/why not.

   (b) Another colleague suggest that you instead run the following regression:

   $$Y_{it} = \lambda_0 + \lambda_1 Sweden_i + \lambda_2 AftMarch_t + \lambda_3 AftMarch_t \times Sweden_i + u_{it},$$

   where $Sweden_i$ is a dummy which is 1 if country i is Sweden and 0 otherwise. $AftMarch_t$ is a dummy which is 1 when the observation is recorded in March or later. What kind of research design is this? Interpret the coefficients in this regression. Under which conditions can this model be used to measure the effect of lockdowns? Discuss.

   (c) One of the countries in your dataset is Denmark. Malmø, the third largest city in Sweden, and the Danish capital Copenhagen, are part of the same metropolitan region with many people commuting across the border. Denmark chose to lock down mid-March. Is this a problem for the research design in question 3.b? (Hint: Discuss SUTVA)

(d) In a paper in the Econometrics Journal, Cho, S. W. (2020) seeks to estimate the effect of lockdowns.[1] Using weekly data on infection rates for Sweden along with Finland, Greece, Norway, Denmark, and Estonia. All countries beside Sweden implemented lockdowns at some point in the sample period. Cho estimates the following model:

$$Y_{it} = \eta_i + \nu_t + \theta L_{it} + \varepsilon_{it},$$

where Y_{it} denotes infection cases per million inhabitants in country i at time t, η_i and ν_t are country and time fixed effects respectively, and L_{it} is a binary treatment dummy which equals 1 if country i was locked down at time t. The treatment dummy will equal 1 for at least one period for all countries except Sweden. How does this model differ from the model in question 4.b? What does an estimate of θ measure? Under what conditions can θ be given a causal interpretation.

Table 3: Results from Cho (2020)

	Infection rate (Y_{it})
Lockdown (L_{it})	-751.1
S.E.	(90.75)
N	600
R^2	0.773

(e) How can the equation in 3.d be estimated? Why is there no constant?

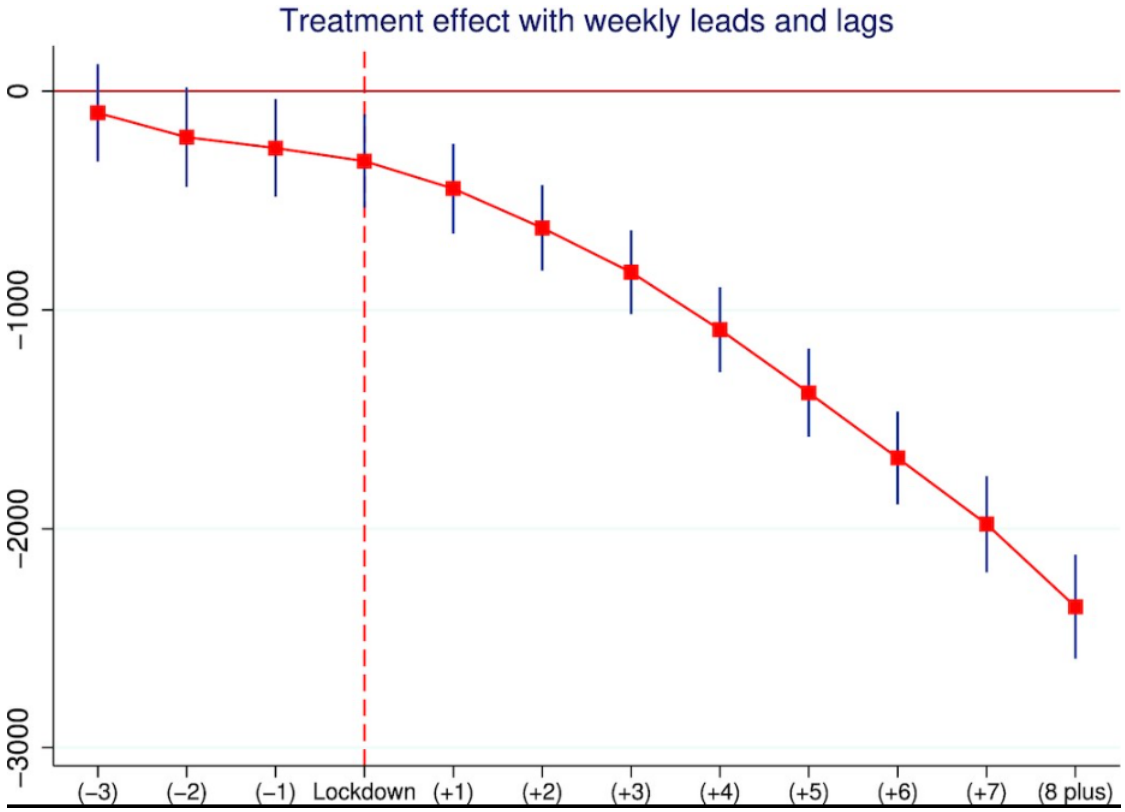
(f) The results of this estimation is presented in Table 3. Interpret the result. Is the R^2 in line with expectations and is it important? Is the parameter on L_{it} significant?

(g) Cho writes "*Although most [...] applications customarily cluster standard errors at the aggregate level, there are too few units of observation to follow clustering methods. As such, I report heteroscedasticity robust standard errors.*". In light of this quote, reflect on your answer to the previous question.

(h) Cho presents the figure displayed in Figure 1. What does this figure show? How would you assess the research design in light of this figure? Discuss.

[1] Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. The Econometrics Journal, 23(3), 323-344.

Figure 1: Anticipatory and time-varying effects



Notes: Horizontal axis shows weekly leads and lags from the day of lockdown. Vertical bands are 95 percent confidence bands. Source: Econom J, Volume 23, Issue 3, September 2020, Pages 323–344, https://doi.org/10.1093/ectj/utaa025