



Stockholm
University

Department of Economics

Course name: Empirical Methods in Economics 2
Course code: EC2404
Type of exam: Main
Examiner: Peter Skogman Thoursie
Number of credits: 7,5 hp
Date of exam: October 31, 2019
Examination time: 15:00-18:00
Aids: No aids are allowed.

Write your identification number on each answer sheet (the number stated in the upper right hand corner on your exam cover).

Start each new question on a new answer sheet.

Explain notions/concepts and symbols. If you think that a question is vaguely formulated, specify the conditions used for solving it. Only legible exams will be marked.

The exam consists of 5 questions. 100 points in total. Each question is worth 20 points, 100 points in total. For the grade E 45 points are required, for D 50 points, C 60 points, B 75 points and A 90 points.

Note Question 5 is the credit question

Your results will be made available on your Ladok account (www.student.ladok.se) within 15 working days from the date of the examination.

Good luck!

Question 1 – Multiple choice (20 points, 4 points each)

- 1) In terms of the potential outcome model, $E[Y_{0i}|D_i = 1] = E[Y_{0i}|D_i = 0]$ implies
- A) that there is no effect of treatment
 - B) that treated are equal in a statistical sense with respect to pre-outcome as well as post-outcome variables
 - C) there is no selection bias
 - D) that we can't construct the counterfactual state
- 2) In a multiple regression, $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$, both X_1 and X_2 are determinants of the dependent variable. The regression slope coefficient in a regression of X_2 on X_1 is γ . When omitting X_2 from the regression, then there will be omitted variable bias for $\hat{\beta}_1$
- A) only if both β_2 and γ are zero
 - B) if the product between β_2 and γ is zero
 - C) if $\beta_2 = -1$ and $\gamma = 1$
 - D) if X_2 is a dummy variable
- 3) In an IV analysis, the exclusion restriction is required in order to estimate the causal effect from the
- A) first stage regression
 - B) reduced form outcome equation
 - C) main equation of interest (i.e., the second stage regression)
 - D) all regressions in A, B and C
- 4) Presence of compliers in an IV analysis, implies
- A) we cannot estimate the causal effect
 - B) that we can only estimate the causal effect of those who take treatment
 - C) that we can only estimate the causal effect of those who react to the instrument
 - D) that we can only estimate the causal effect from the reduced form equation
- 5) In a large-scale randomized experiment, the reason for controlling for pre-determined characteristics is
- A) to solve for a potential omitted variable bias problem
 - B) make the exclusion restriction assumption more likely to hold
 - C) is to detect mechanisms
 - D) increase the precision of the estimated treatment effect

Question 2 – Multiple choice (20 points, 4 points each)

1) The key identifying assumption for causality in the IV strategy when the instrument, Z , is as good as randomized controlling for X , can be written as

- A) $E(u|Z, X) = E(u|X)$
- B) $E(u|Z, X) = E(u|Z)$
- C) $E(u|Z) = 0$
- D) $E(u|Z, X) = 0$

2) One key identifying assumption for causality in the Regression discontinuity method is that

- A) the forcing variable, X , is uncorrelated with treatment, D
- B) potential outcomes are continuous around the cutoff values
- C) $D = 1$ and $D = 0$ are as good as randomly distributed
- D) that the relationship between the forcing variable and the outcome is correctly specified

3) When formulating the RD regression: $Y = \alpha + \rho D + \gamma_1(X - X_0) + \gamma_2(X - X_0) \times D + u$, where X_0 is the cut-off value, ρ measures

- A) the effect of D on Y at $(X - X_0)$ units above the cut-off
- B) the marginal effect of D when $X = 0$
- C) the discrete jump in Y at the cut-off value
- D) the marginal effect of D evaluated at then mean of X

4) Estimating the DD model $Y_{st} = \alpha_s + \beta After_t + \delta After_t \times D_s + \gamma X_{st} + u_{st}$, you need to assume that

- A) $D = 1$ and $D = 0$ are as good as randomized once you control for X_{st}
- B) $D = 1$ and $D = 0$ are as good as randomized once you control for $After_t$
- C) $D = 1$ and $D = 0$ are as good as randomized once you control for X_{st} and $After_t$
- D) $D = 1$ and $D = 0$ have the same potential outcomes in absence of treatment once α_s , $After_t$ and X_{st} are controlled for

5) When estimating the panel data model $Y_{it} = \alpha_i + \beta X_{it} + u_{it}$

- A) you need to cluster standard errors at the time level
- B) you need to use heteroskedasticity robust standard errors to account for any form of heteroscedasticity
- C) you need to cluster standard errors at the individual level
- D) you need to cluster standard errors and use heteroskedasticity robust standard errors at the same time

Question 3 – Difference-in-differences (20 points)

Economic models predict that the marginal tax rate have ambiguous effects on labour supply. A lower marginal tax rate implies higher income, which reduces labour supply (if leisure is a normal good) through the income effect. However, the price of leisure increases relative to work implying increased labour supply through the substitution effect.

Say that half of the municipalities in Sweden lowered the marginal tax rate in 2010.

Your task is to write a research outline on how you will use this reform to estimate the casual effect of the marginal tax rate on labour supply.

There are couple of issues that need to be incorporated in this outline.

- (i) What is the population of interest and what type of data that is required
- (ii) Equations that are supposed to be estimated within the project should be clearly and correctly specified and explained
- (iii) Explain how you would interpret the estimated effects from the specified regression equations
- (iv) How to evaluate the key identifying assumption
- (v) Individuals might learn about the tax system with a lag. Thus, the research design needs to explain how to deal with the issues that potential effects might show up gradually over years
- (vi) Labor supply trends might exhibit differential trends across municipalities. The outline needs to econometrically address this issue

The outline should be maximum one and half pages. Write short and precise!

**Question 4 – IV (20 points)**

Some scholars have observed that there are some countries where there are preferences for sons. For example, Amartya Sen has observed that this has led to fewer girls being born.

For simplicity, let's assume that the possibility to know the sex before birth is ruled out.

Your task is to write a research outline on how you will estimate the effect of fertility on labour supply using the sex of the first-born child as an instrument to fertility.

There are couple of issues that need to be incorporated in this outline.

- (i) What is the population of interest and what type of data that is required
- (ii) Discuss the key identifying assumptions (relevance, randomization and exclusion restriction)
- (iii) All three equations part of the IV set-up should be clearly and correctly specified
- (iv) How the slope regression coefficient should be interpreted in each of these three equations (note that there might be heterogeneous effects)

The outline should be maximum one and half pages. Write short and precise!

Question 5 – credit question. Acemoglu & Angrist (2001) paper (20 points)

This is the abstract from the Acemoglu & Angrist (2001) paper:

“The Americans with Disabilities Act (ADA) requires employers to accommodate disabled workers and outlaws discrimination against the disabled in hiring, firing, and pay. Although the ADA was meant to increase the employment of the disabled, the net theoretical effects are ambiguous. For men of all working ages and women under 40, Current Population Survey data show a sharp drop in the employment of disabled workers after the ADA went into effect. Although the number of disabled individuals receiving disability transfers increased at the same time, the decline in employment of the disabled does not appear to be explained by increasing transfers alone, leaving the ADA as a likely cause. Consistent with this view, the effects of the ADA appear larger in medium-size firms, possibly because small firms were exempt from the ADA. The effects are also larger in states with more ADA-related discrimination charges.”

Describe how they have econometrically reached the main conclusion that the ADA seems to have a negative effect on employment of the disabled. Especially, the following issues must be included and explained intuitively as well as using equation notations:

- (i) The main strategy used?
- (ii) What is the key identifying assumption for estimating the causal effect of the ADA?
- (iii) How do they econometrically investigate if this assumption is valid?

Write maximum 1 ½ A4-page for the answers!