Student: _____ **Date:** _____	**Instructor:** Richeng Piao **Course:** ECON 2560 - Applied Econometrics	**Assignment:** Practice Problem Set 13

I hereby declare and affirm that I will not redistribute the practice question set. I understand and acknowledge that the practice questions provided to me are intended solely for personal use and reference. I will not share, copy, reproduce, distribute, or make the practice question set available to any third parties without explicit authorization from the rightful owner or the authorized distributor. I respect the intellectual property rights and confidentiality associated with the practice question set and will adhere to the terms and conditions stated.

Signature _____ Date _____

1. A researcher studying the effects of a new fertilizer on crop yields plans to carry out an experiment in which different amounts of the fertilizer are applied to 100 different 1-acre parcels of land. There will be four treatment levels. Treatment level 1 is no fertilizer, treatment level 2 is 50% of the manufacturer's recommended amount of fertilizer, treatment level 3 is 100%, and treatment level 4 is 150%. The researcher plans to apply treatment level 1 to the first 25 parcels of land, treatment level 2 to the second 25 parcels, and so forth.

   Can you suggest a better way to assign treatment levels?

   - **A.** Assign the fertilizer treatment level randomly to each parcel.
   - **B.** Assign 150% of the manufacturer's recommended amount of fertilizer level to each parcel.
   - **C.** Assign 100% of the manufacturer's recommended amount of fertilizer level to each parcel.
   - **D.** Assign 50% of the manufacturer's recommended amount of fertilizer level to each parcel.

   Why is your proposal better than the researcher's method?

   - **A.** The researcher's method may be flawed because it does not assign equal amounts of fertilizer to all parcels.
   - **B.** The researcher's method may be flawed because the different groups of parcels might differ systematically.
   - **C.** The researcher's method may be flawed because it involves too few treatment levels.
   - **D.** The researcher's method is indeed the optimal design.

   Answers A. Assign the fertilizer treatment level randomly to each parcel.

   B. The researcher's method may be flawed because the different groups of parcels might differ systematically.

   ID: Review Concept 13.1

2. A clinical trial is carried out for a new cholesterol-lowering drug. The drug is given to 500 patients and a placebo given to another 500 patients, using random assignment of the patients.

   How would you estimate the treatment effect of the drug?

   - **A.** It could be estimated by observing the average cholesterol level of all individuals from both groups over a period of time.
   - **B.** It could be estimated as the difference in average cholesterol levels for the treated group and the untreated (control) group.
   - **C.** It could be estimated by comparing the cholesterol levels of two random individuals in the treated group.
   - **D.** It could be estimated by comparing the cholesterol levels of one randomly chosen individual from each group.

   Suppose that you had data on the weight, age, and gender of each patient.

   Could you use these data to improve your estimate? Explain.

   - **A.** The regression may not produce a more accurate estimate because the additional information is not related to cholesterol.
   - **B.** A regression with these additional regressors may produce a less accurate estimate because it considers these additional factors th
   - **C.** Data on the weight, age, and gender of each patient could not be used to improve the estimate since they will lead to multicollineari
   - **D.** Data on the weight, age, and gender of each patient could be used to improve the estimate using the differences estimator with the

   Suppose that you had data on the cholesterol levels of each patient before he or she entered the experiment.

   Could you use these data to improve your estimate? Explain.

   - **A.** The differences-in-differences estimator, which controls for individual-specific determinants of cholesterol levels that are constant o
   - **B.** The data will not accurately demonstrate the person's genetic predisposition to high cholesterol.
   - **C.** The differences-in-differences estimator, which controls for individual-specific determinants of cholesterol levels that are constant o
   - **D.** The use of the additional data will result in a noncompliance with the experiment protocol.

   Answers B.
   It could be estimated as the difference in average cholesterol levels for the treated group and the untreated (control) group.

   D.
   Data on the weight, age, and gender of each patient could be used to improve the estimate using the differences estimator with these additional regressors.

   C.
   The differences-in-differences estimator, which controls for individual-specific determinants of cholesterol levels that are constant over the sample period, can be used.

   ID: Review Concept 13.2

3. Suppose that, in a randomized controlled experiment of the effect of an SAT preparatory course on SAT scores, the following results are reported:

	Treatment Group	Control Group
Average SAT score	1,241	1,208
Standard deviation of SAT score	93.9	97.2
Number of men	51	49
Number of women	49	51

Estimate the average treatment effect on test scores.

Average treatment effect = [ ] points. (*Enter your response as an integer.*)

There would be nonrandom assignment if men (or women) had different probabilities of being assigned to the treatment group and to the control group. Let p_{Men} denote the probability that a male is assigned to the treatment group. Random assignment to a group means $p_{Men} = 0.5$.

Calculate the *t*-statistic for random assignment of a male to a group under the null hypothesis of $p_{Men} = 0.5$ at the 10% level of significance.

t-statistic = [ ]. (*Round your response to two decimal places.*)

Using this calculation, there (1) _____ evidence of nonrandom assignment.

(1) ○ is not
○ cannot be determined
○ is

Answers 33

0.200

(1) is not

ID: Exercise 13.3

---

4. In the context of a controlled experiment, consider the simple linear regression formulation $Y_i = \beta_0 + \beta_1 X_i + \mu_i$. Let the Y_i be the outcome, X_i the treatment level when the treatment is binary, and μ_1 contain all the additional determinants of the outcome.

Then calling $\hat{\beta}_1$ a differences estimator:

○ **A.** does not make sense, since neither *Y* nor *X* are in differences.
○ **B.** and $\hat{\beta}_0$ the level estimator is standard terminology in randomized controlled experiments.
○ **C.** is not quite accurate since it is actually the derivative of *Y* on *X*.
○ **D.** makes sense since it is the difference between the sample average outcome of the treatment group and the sample average outcor

Answer: D.

makes sense since it is the difference between the sample average outcome of the treatment group and the sample average outcome of the control group.

ID: Test A Ex 13.1.1

---

5. Experimental data are often:

- **A.** panel data.
- **B.** binary data, in that the subject either does or does not respond to the treatment.
- **C.** time series data.
- **D.** observational data.

Answer: A. panel data.

ID: Test A Ex 13.1.2

---

6. A pharmaceutical company wants to test the effectiveness of a newly synthesized drug against the HIV virus. For this, the company selects a group of 330 HIV patients randomly. It then administers a placebo to a control group, anti-retroviral treatment (*ART*) to one test group, and the newly synthesised drug (*Cure*) to another test group for 17 months. The patients are equally assigned to each of these three groups randomly. At the end of 17 months, it measures the effectiveness of the newly synthesized drug and the anti-retroviral treatment against the control group by regressing the binary variables *Cure* and *ART* on their effectiveness in delaying the replication of the HIV virus in a patient (*CD4*). *Cure* = 1 if the newly synthesized drug is administered, and *Cure* = 0 otherwise; and, *ART* = 1 if the anti-retroviral treatment is administered and *ART* = 0 otherwise. *CD4* is measured by the CD4 cell count per cubic millimetre of blood in an HIV patient. CD4 cells are white blood cells that define the immune system in a human body and are inversely related to the rate of replication of the HIV virus in a patient's body. Only one type of treatment is administered to one group throughout the course of the experiment. The company estimates the following regression equation:

$$\widehat{CD4} = 350 + 75ART + 86Cure.$$
$$(20.34)\ (21.72)\ \ (23.91)$$

Standard errors are given in parentheses.

The average causal effect of the newly synthesized drug is [ ] $CD4/mm^3$.

The company wants to test whether the effect of the newly synthesized drug in delaying the replication of the HIV virus in a patient is significant or not.

The test statistic associated with the test the company wants to conduct is [ ].

(*Round your answer to two decimal places.*)

At the 1% significance level, the company (1) _____ the hypothesis that the effect of the newly synthesized drug in delaying the replication of the HIV virus in a patient is not significant.

The difference in the average causal effects of the anti-retroviral treatment and the newly synthesized drug is [ ] $CD4/mm^3$.

From the experiment, the pharmaceutical company concludes that the newly synthesized drug is (2) _____ the anti-retroviral treatment for HIV.

(1) ○ fails to reject ○ rejects

(2) ○ more effecive than ○ less effective than

Answers 86

3.6

(1) rejects

11

(2) more effecive than

ID: Concept Exercise 13.1.1

---

7. Suppose we have a multiple regression equation as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{1i} + \ldots + \beta_{1+r} W_{ri} + u_i,\ i = 1, \ldots, n,$$

where X_i is a binary regressor, $W_{1i} \ldots W_{ri}$ are the control regressors, and β_0, β_1,.., β_{1+r} are the intercept and the slope coefficients on the binary regressor and the control regressors, respectively.

Which of the following statements correctly describe the features of a difference estimator with multiple regressors?

- ☐ **A.** If W_i is a pretreatment characteristic and X_i is randomly assigned, then X_i is independent of u_i and W_i.
- ☐ **B.** The control regressors should include experimental outcomes; i.e., X_i is randomly assigned, given an experimental outcome.
- ☐ **C.** If W helps to explain the variation in Y, then including W reduces the standard error of the regression and, typically, the standard er
- ☐ **D.** The coefficients on the control variables do not have causal interpretations.

Randomization in which the probability of assignment to the treatment group depends on one or more (1) _____ is called randomization based on covariates.

(1)
- ○ observable variables
- ○ experimental outcomes

Answers A. If W_i is a pretreatment characteristic and X_i is randomly assigned, then X_i is independent of u_i and W_i., C. If W helps to explain the variation in Y, then including W reduces the standard error of the regression and, typically, the standard error of $\hat{\beta}_1$.
, D. The coefficients on the control variables do not have causal interpretations.

(1) observable variables

ID: Concept Exercise 13.1.2

---

8. Explain whether experimental effects (like the Hawthorne effect) might be important in each of the following experiments: a crop yield experiment in which different amounts of fertilizer are applied to 100 different 1-acre parcels of land; a clinical trial for a new cholesterol-lowering drug using random assignment of 1,000 patients, and the STAR experiment in which elementary students were randomly assigned to different-sized classes in the state of Tennessee in the late 1980s.

- ○ **A.** The Hawthorne effect is unlikely to pose a problem for the fertilizer crop yield study or the STAR experiment, but is likely to be pres
- ○ **B.** The Hawthorne effect is likely to be present in all three experiments.
- ○ **C.** The Hawthorne effect is unlikely to pose a problem for the fertilizer crop yield study, but is likely to be present in both the cholestero
- ○ **D.** The Hawthorne effect is likely to be present in the fertilizer crop yield study and the cholesterol study, but is unlikely to be present in

Answer: C.
The Hawthorne effect is unlikely to pose a problem for the fertilizer crop yield study, but is likely to be present in both the cholesterol study and the STAR experiment.

ID: Review Concept 13.4

---

9. Consider a study to evaluate the effect on college student grades of dorm room Internet connections. In a large dorm, half the rooms are randomly wired for high-speed Internet connections (the treatment group), and final course grades are collected for all residents.

Which of the following pose threats to internal validity?

(i.) Midway through the year, all the male athletes move into a fraternity and drop out of the study. (Their final grades are not observed.)

(ii.) Engineering students assigned to the control group put together a local area network so that they can share a private wireless Internet connection that they pay for jointly.

(iii.) The art majors in the treatment group never learn how to access their Internet accounts.

(iv.) The economics majors in the treatment group provide access to their Internet connection to those in the control group, for a fee.

- **A.** Only (ii) and (iv) due to partial compliance.
- **B.** Only (i), (ii), and (iv) pose threats to internal validity.
- **C.** Only (i) due to attrition and (iii) due to failure to follow protocol.
- **D.** All of the scenarios above pose threats to internal validity.

Answer: B. Only (i), (ii), and (iv) pose threats to internal validity.

ID: Exercise 13.5

---

10. Small sample sizes in an experiment:

- **A.** may pose a problem because the assumption that errors are normally distributed is dubious for experimental data.
- **B.** bias the estimators of the causal effect.
- **C.** may affect confidence intervals but not hypothesis tests.
- **D.** do not raise threats to the validity of confidence intervals as long as heteroskedasticity-robust standard errors are used.

Answer: A. may pose a problem because the assumption that errors are normally distributed is dubious for experimental data.

ID: Test A Ex 13.2.3

---

11. The following is *not* a threat to external validity:

- **A.** the treatment being studied is not representative of the treatment that would be implemented more broadly.
- **B.** the experimental sample is not representative of the population of interest.
- **C.** partial compliance with the treatment protocol.
- **D.** experimental participants are volunteers.

Answer: C. partial compliance with the treatment protocol.

ID: Test B Ex 13.2.1

---

12. The following estimation method should *not* be used to test for randomization when X_i is binary:

- **A.** Logit.
- **B.** Linear probability model (OLS) with heteroskedasticity-robust standard errors.
- **C.** Linear probability model (OLS) with homoskedasticity-only standard errors.
- **D.** Probit.

Answer: C. Linear probability model (OLS) with homoskedasticity-only standard errors.

ID: Test B Ex 13.2.2

---

13. Testing for the random receipt of treatment:

- **A.** entails testing the hypothesis that the coefficients on $W_{1i},...,W_{ri}$ are zero in a regression of X_i on $W_{1i},...,W_r$.
- **B.** is not possible, in general.
- **C.** entails testing the hypothesis that the coefficients on $W_{1i},...,W_{ri}$ are nonzero in a regression of X_i on $W_{1i},...,W_r$.
- **D.** is not meaningful since the LHS variable is binary.

Answer: A. entails testing the hypothesis that the coefficients on $W_{1i},...,W_{ri}$ are zero in a regression of X_i on $W_{1i},...,W_r$.

ID: Test B Ex 13.2.3

---

14. Which of the following statements does not describe a possible threat to the internal validity of randomized control experiments?

- **A.** There is a threat to the internal validity if the treatment is not assigned randomly, but instead is based in part on the characteristics
- **B.** There is a threat to the internal validity if subjects drop out of the study after being randomly assigned to the treatment or the contr
- **C.** There is a threat to the internal validity if individuals fail to follow completely the randomized treatment protocol.
- **D.** There is a threat to the internal validity if the population studied and the population of interest are not sufficiently similar to justify g

Which of the following statements describe possible threats to the external validity of randomized control experiments? (*Check all that apply*.)

- ☐ **A.** There is a threat to the external validity if the policy or program of interest is not sufficiently similar to the program studied to permi
- ☐ **B.** There is a threat to the external validity if the sample size is small as this raises threats to the validity of confidence intervals and h
- ☐ **C.** There is a threat to the external validity if turning a small, temporary experimental program into a widespread, permanent program
- ☐ **D.** There is a threat to the external validity if in experiments with human subjects, merely because the subjects are in an experiment c

Answers D.
There is a threat to the internal validity if the population studied and the population of interest are not sufficiently similar to justify generalizing the results.

A.
There is a threat to the external validity if the policy or program of interest is not sufficiently similar to the program studied to permit generalizing the results.
, C.
There is a threat to the external validity if turning a small, temporary experimental program into a widespread, permanent program might change the economic environment sufficiently that the results from the experiment cannot be generalized.

ID: Concept Exercise 13.2.1

---

15. Researchers studying the STAR data report anecdotal evidence that school principals were pressured by some parents to place their children in the small classes. Suppose that some principals succumbed to this pressure and transferred some children into the small classes.

How would such transfers compromise the internal validity of the study?

- **A.** The internal validity is compromised if the students who were transferred to the small classes differed systematically from the othe
- **B.** The internal validity is compromised if the transferred students tended to have higher incomes and more learning opportunities out
- **C.** The internal validity is compromised because the transfer compromises the control group of the experiment.
- **D.** Both (a) and (b) are true.

Suppose that you had data on the original random assignment of each student before the principal's intervention.

How could you use this information to restore the internal validity of the study?

- **A.** It could be used as an instrument in a regression to restore internal validity because the original random assignment is correlated
- **B.** It could be used as an instrument in a regression to restore internal validity because the original random assignment is not exogen
- **C.** It could be used as an instrument in a regression to restore internal validity because the original random assignment is exogenous
- **D.** It could not be used as an instrument in a regression to restore internal validity because the original random assignment is exogen

Answers D. Both (a) and (b) are true.

C.
It could be used as an instrument in a regression to restore internal validity because the original random assignment is exogenous or uncorrelated with the regression error and is relevant or correlated with the actual assignment.

ID: Review Concept 13.3

16. **Table 1: Differences Estimates of Effect on Standardized Test Scores of Class Size Treatment Group**

	Grade			
Regressor	K	1	2	3
Small class	13.06** (2.43)	29.34** (2.85)	19.76** (2.75)	15.61** (2.33)
Regular size with aide	0.31 (2.27)	11.96 (2.65)	3.48 (2.54)	-0.29 (2.27)
Intercept	918.04** (1.63)	1039.39** (1.78)	1157.81** (1.82)	1228.51** (1.68)
Number of observations	5786	6379	6049	5967

Note: The dependent variable is the student's combined score on the math and reading portions of an achievement test. Standard errors are given in parentheses under the coefficients. ** The individual coefficient is statistically significant at the 1% significance level using a two-sided test.

Using the results in the table above, calculate the following for each grade: an estimate of the small class treatment effect, relative to the regular class; its standard error; and its 95% confidence interval. (For this exercise, ignore the results for regular classes with aides.)

For students in kindergarten:

The estimated small class treatment effect relative to being in a regular class is an increase of _____ points on the test. (*Round your response to two decimal places.*)

Standard error = _____. (*Round your response to two decimal places.*)

The 95% confidence interval = (_____, _____). (*Round your responses to two decimal places.*)

For students in grade 1:

The estimated small class treatment effect relative to being in a regular class is an increase of _____ points on the test. (*Round your response to two decimal places.*)

Standard error = _____. (*Round your response to two decimal places.*)

The 95% confidence interval = (_____, _____). (*Round your responses to two decimal places.*)

For students in grade 2:

The estimated small class treatment effect relative to being in a regular class is an increase of _____ points on the test. (*Round your response to two decimal places.*)

Standard error = _____. (*Round your response to two decimal places.*)

The 95% confidence interval = (_____, _____). (*Round your responses to two decimal places.*)

For students in grade 3:

The estimated small class treatment effect relative to being in a regular class is an increase of _____ points on the test. (*Round your response to two decimal places.*)

Standard error = _____. (*Round your response to two decimal places.*)

The 95% confidence interval = (_____, _____). (*Round your responses to two decimal places.*)

Answers 13.06

2.43

8.297

17.823

29.34

2.85

23.754

34.926

19.76

2.75

14.370

25.150

15.61

2.33

11.043

20.177

ID: Exercise 13.1

---

17. A randomised control experiment was conducted to estimate the effectiveness of a new FDA approved memory-enhancing drug on the test scores of college students. From a group of students with similar IQ levels, 400 students were randomly selected for the experiment and were randomly assigned to a treatment and a control group in equal numbers. The treatment group was asked to consume the FDA approved memory-enhancing drug for a certain period of time (*Drug* = 1) and the control group was given a placebo (*Drug* = 0). The objective was to check for any significant increase in the first semester average test scores (*Testscore*) of the treatment group due to the consumption of the drug. The estimated regression is given by:

$$\widehat{Testscore} = 35.19 + 8.45Drug.$$
$$(1.09)\ \ (5.23)$$

Standard errors are given in parentheses.

The 95% confidence interval for the difference in the average test scores between the treatment and the control group will be ([ ] , [ ]).

(*Round your answers to two decimal places. Enter a minus sign if your answer is negative.*)

Let β_1 denote the slope coefficient on *Drug*.

Based on the calculated confidence interval, we can say that at the 5% significance level, we (1) _____ the hypothesis that $\beta_1 = 0$.

(1) ○ fail to reject
○ reject

Answers – 1.80

18.70

(1) fail to reject

ID: Concept Exercise 13.3.1

---

18. Suppose that you have panel data from an experiment with $T = 2$ periods (so $t = 1, 2$). Consider the panel data regression model with fixed individual and time effects and individual characteristics W_i that do not change over time, such as gender. Let the treatment be binary, so that $X_{it} = 1$ for $t = 2$ for the individuals in the treatment group and let $X_{it} = 0$ otherwise. Consider the population regression model:

$$Y_{it} = \alpha_i + \beta_1 X_{it} + \beta_2 \left(D_t \times W_i\right) + \beta_0 D_t + v_{it},$$

where α_i are individual fixed effects, D_t is the binary variable that equals 1 if $t = 2$ and equals 0 if $t = 1$, $\left(D_t \times W_i\right)$ is the product of D_t and W_i, and the α's and β's are unknown coefficients. Let $\Delta Y_i = Y_{i2} - Y_{i1}$.

Write the above regression equation as a multiple regression equation with ΔY_i as the dependent variable, which can be used for the differences-in-differences estimation (that is, in the form of Equation 13.6 from the text).

$\Delta Y_i =$ [ ]. *(Properly format your expression using the tools in the palette. Hover over tools to see keyboard shortcuts. E.g., a subscript can be created with the _ character.)*

Answer: $\beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$

ID: Exercise 13.7

---

19. Suppose that you have panel data from an experiment with $T = 2$ periods (so $t = 1, 2$). Consider the panel data regression model with fixed individual and time effects. Let the treatment be binary, so that $X_{it} = 1$ for $t = 2$ for the individuals in the treatment group and let $X_{it} = 0$ otherwise. Consider the population regression model:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 G_i + \beta_3 D_t + u_{it},$$

where $G_i = 1$ if the individual is in the treatment group, $G_i = 0$ if the individual is in the control group, and D_t is the binary variable that equals 1 if $t = 2$ and 0 if $t = 1$.

Calculate the differences-in-differences estimator, $\widehat{\beta}_1^{diffs-in-diffs}$, which is the average change in Y for those in the treatment group minus the average change in Y for those in the control group, or $\widehat{\beta}_1^{diffs-in-diffs} = \Delta\bar{Y}^{treatment} - \Delta\bar{Y}^{control}$.

$\widehat{\beta}_1^{diffs-in-diffs} =$ [ ]. *(Properly format your expression using the tools in the palette. Hover over tools to see keyboard shortcuts. E.g., a subscript can be created with the _ character.)*

Answer: $\widehat{\beta}_1$

ID: Exercise 13.8

---

20. Quasi-experiments:

- **A.** are not the same as experiments, and lessons learned from the use of the latter can therefore not be applied to them.
- **B.** provide a bridge between the econometric analysis of observational data sets and the statistical ideal of a true randomized control
- **C.** use the same methods as studied in earlier chapters of the book, and hence the interpretation of these methods is the same.
- **D.** most often use difference-in-difference estimators, which are quite different from OLS and instrumental variables methods studied

Answer: B.
provide a bridge between the econometric analysis of observational data sets and the statistical ideal of a true randomized controlled experiment.

ID: Test A Ex 13.4.4

---

21. In a sharp regression discontinuity design:

- **A.** receipt of treatment is entirely determined by whether W exceeds the threshold.
- **B.** X_i will in general be correlated with μ_i.
- **C.** crossing the threshold influences receipt of the treatment but is not the sole determinant.
- **D.** the population regression line must be linear above and below the threshold.

Answer: A. receipt of treatment is entirely determined by whether W exceeds the threshold.

ID: Test A Ex 13.4.5

---

22. A repeated cross-sectional data set is:

- **A.** what Card and Krueger used in their study of the effect of minimum wages on teenage employment.
- **B.** a collection of cross-sectional data sets, where each cross-sectional data set corresponds to a different time period.
- **C.** time series.
- **D.** the same as a balanced panel data set.

Answer: B. a collection of cross-sectional data sets, where each cross-sectional data set corresponds to a different time period.

ID: Test B Ex 13.4.4

---

23. Nuclear weapons testing releases large quantities of radioactive isotopes into the atmosphere, some of which are incorporated into biological tissues. The release stopped after the Partial Nuclear Test Ban Treaty in 1963, which prohibited atmospheric nuclear tests. Although, some countries that did not sign the treaty continued atmospheric nuclear weapons testing. A researcher wants to test the effect of the Partial Nuclear Test Ban Treaty, on the cancer rate, (*Canrate*, measured as percentage of cancer patients in a country). He uses data from two randomly selected countries A and B from 1942 to 1996, such that country A is a signatory and country B is a non-signatory of the Partial Nuclear Test Ban Treaty. Country B continued atmospheric nuclear testing after 1963 until 1996. The researcher uses the data from country A as the treatment group and the data from country B as the control group.

The average cancer rates calculated using the data from 1942 to 1963 in country A ($\overline{Canrate}^{A,\ before}$) and country B ($\overline{Canrate}^{B,\ before}$) are 11.86% and 16.17%, respectively. The average cancer rates calculated using the data from 1964 to 1985 in countries A ($\overline{Canrate}^{A,\ after}$) and B ($\overline{Canrate}^{B,\ after}$) are 4.23% and 9.58%, respectively.

The differences-in-differences estimator of the causal effect of the treaty is [ ] %.

(*Round your answer to two decimal places. Enter a minus sign if your answer is negative.*)

The researcher isolates data for two specific years, 1957 and 1975, for countries A and B and estimates the following regression equation:

$$\widehat{Canrate}_{it} = 0.11 - 0.005X_{it} - 0.024G_i - 0.036D_t,\ (i = \text{A, B and } t = 1957, 1975)$$
$$(0.013)\ (0.007)\ \ (0.0012)\ \ \ (0.0031)$$

where G is a binary variable that takes the value 1 if the data is from the treatment group and 0 otherwise, D is a binary variable that takes the value 1 if the data is for 1975 and 0 if the data is for 1957, and $X = (G \times D)$. Standard errors are given in parentheses.

The differences-in-differences estimator of the causal effect of the treaty is [ ].

(*Round your answer to three decimal places. Enter a minus sign if your answer is negative.*)

Answers -1.04

-0.005

ID: Concept Exercise 13.4.1

24. Which of the following is not a condition that must be satisfied such that the instrumental variable used in a quasi-experiment is a valid instrument for the treatment actually received?

- **A.** The instrument should influence the receipt of the treatment.
- **B.** The instrument should be "as if" randomly assigned.
- **C.** There should be complete compliance with the treatment protocol.
- **D.** The data should be available both on the instrument and on the treatment actually received.

Studies that exploit a discontinuity in the probability of receiving treatment at (1) _____ value are called regression discontinuity designs.

Which of the following statements are true about regression discontinuity designs? (*Check all that apply*.)

- **A.** In a sharp regression discontinuity design, receipt of treatment is entirely determined by whether the control variable exceeds the t
- **B.** In a sharp regression discontinuity design, the slope estimator can be estimated by instrumental variables estimation given that th
- **C.** In a fuzzy regression discontinuity design, the jump in regressand value at the threshold equals the average treatment effect for th
- **D.** In a fuzzy regression discontinuity design, crossing the threshold influences receipt of the treatment but is not the sole determinan

(1)
- the median
- the mean
- a threshold

Answers C. There should be complete compliance with the treatment protocol.

(1) a threshold

A.
In a sharp regression discontinuity design, receipt of treatment is entirely determined by whether the control variable exceeds the threshold or not.
, D.
In a fuzzy regression discontinuity design, crossing the threshold influences receipt of the treatment but is not the sole determinant of whether the treatment will be received.

ID: Concept Exercise 13.4.2

---

25. Threats to internal validity of quasi-experiments include:

- **A.** attrition.
- **B.** failure to follow the treatment protocol.
- **C.** failure of randomization.
- **D.** all of the above with some modifications from true randomized controlled experiments.

Answer: D. all of the above with some modifications from true randomized controlled experiments.

ID: Test B Ex 13.5.5

---

26. Consider the quasi-experiment described in Section 13.4 involving the draft lottery, military service, and civilian earnings.

Explain why there might be heterogeneous effects of military service on civilian earnings; that is, explain why β_{1i} in equation $Y_i = \beta_{0i} + \beta_{1i}X_i + u_i$ depends on *i*.

- **A.** The military accepts only those individuals who meet specific physical requirements.
- **B.** Individuals joining the military may have been unsuccessful in the private labor market.
- **C.** Veterans may have acquired different skills while in the military.
- **D.** Both (a) and (b) are true.

Explain why there might be heterogeneous effects of the lottery outcome on the probability of military service; that is, explain why π_{1i} in equation $X_i = \pi_{0i} + \pi_{1i}Z_i + \upsilon_i$ depends on *i*.

- **A.** Actual entry into the military was subject to certain exemptions, such as medical study.
- **B.** Actual entry into the military was controlled by random lottery that determined draft eligibility.
- **C.** Actual entry into the military was based on certain physical abilities.
- **D.** Both (a) and (c) are correct.

If there are heterogeneous responses of the sort described above, what behavioral parameter is being estimated by the TSLS estimator?

- **A.** Average treatment effect.
- **B.** Heterogeneous causal effect.
- **C.** Hawthorne effect.
- **D.** Quasi-experiment effect.

Answers D. Both (a) and (b) are true.

D. Both (a) and (c) are correct.

A. Average treatment effect.

ID: Review Concept 13.5

27. Suppose the population regression equation is given by:

$$Y_i = \beta_{0i} + \beta_{1i} X_i + u_i.$$

Because β_{1i} varies from one individual to the next in the population and the individuals are selected from the population at random, β_{1i} is a random variable that, just like u_i, reflects unobserved variation across individuals.

Which of the following is correct? (*Hint:* Use the definition of the covariance and that, because the actual treatment X_i is random, β_{1i} and X_i are independently distributed.)

- **A.** $\text{cov}\left(\beta_{1i} X_i, X_i\right) = E\left(\beta_{1i} X_i^2\right) - E\left(\beta_{1i} X_i\right) E\left(X_i\right).$
- **B.** $\dfrac{\text{cov}\left\{\beta_{1i} X_i, X_i\right\}}{\sigma_x^2} = \dfrac{\text{cov}\left\{\beta_{0i} + \beta_{1i} X_i, X_i\right\}}{\sigma_x^2} = E\left\{\beta_{1i}\right\} = \widehat{\beta_1}.$
- **C.** $\text{cov}\left(\beta_{1i} X_i, X_i\right) = E\left(\beta_{1i}\right) \sigma_x^2.$
- **D.** All of the above.

Answer: D. All of the above.

ID: Exercise 13.9

28. Suppose that a researcher wants to study the effect of doing an optional online economics problem set (*PS* = 1 if a student does the problem set, and *PS* = 0 otherwise) on the grades students receive on their final economics test (*Grade*, expressed as a percentage). She randomly selects 250 students for her study and then assigns them randomly to the treatment group or the control group in equal numbers. The population under consideration is heterogeneous, so each student has his or her own causal effect, which is the difference in that student's potential test grades when the problem set is solved and when it is not. Suppose that students randomly decide whether or not to do the optional online problem set. The population regression equation can be written as:

$$Grade = \beta_{0i} + \beta_{1i}\, PS_i + u_i,$$

where β_{1i} varies from one student to the next in the population and the students are selected from the population at random. The sample covariance between *PS* and *Grade* was calculated to be 7.48, the variance of *PS* was calculated to be 3.84, and the variance of *Grade* was calculated to be 3.24.

The OLS estimator which is a consistent estimator of the average causal effect of a randomly selected student is [ ].

(*Round your answer to two decimal places.*)

Suppose that the causal effect is estimated by instrumental variables regression of *Grade* on *PS* using ownership of a laptop (*L* = 1 if the student owns a laptop, and *L* = 0 otherwise) as an instrument. The incidence of owning a laptop is random among the students. Assume that the instrument is exogenous and relevant.

The sample covariance between *Grade* and *L* is 6.54, the sample covariance between *PS* and *L* is 4.14, and the sample covariance between *Grade* and *PS* is 7.48.

The TSLS estimator which is a consistent estimator of the local average treatment effect (LATE) is [ ].

(*Round your answer to two decimal places.*)

Suppose the laptops that the students own were randomly assigned by the university to exactly half the student population. Owning a laptop does not have any effect on the decision to do the problem set and consequently, the effect of doing the problem set on the grade the student receives is 1.37. But, not owning a laptop has an effect of −0.66 on the decision to do the problem set and consequently, the effect of doing the problem set on the grade the student receives is 1.07.

The average treatment effect is [ ].

(*Round your answer to two decimal places.*)

Answers 1.95

1.58

1.22

ID: Concept Exercise 13.6.1

29. Which of the following statements are correct in describing the threats to internal and external validity that are or are not faced by quasi-experiments? (*Check all that apply*.)

- ☐ **A.** Just because an instrument is randomly determined or as-if randomly determined in a quasi-experiment does not necessarily mean it is exogenous.
- ☐ **B.** Checking for systematic differences between the treatment and control groups to test for failure of randomization is not possible as some factors in the error term which could be related to the treatment level *X* are unobserved.
- ☐ **C.** Experimental effects such as the Hawthorne effect in a quasi-experiment are similar to such effects in a true experiment, but attrition does not occur in a quasi-experiment as it is not possible for individuals to drop out of such studies.
- ☐ **D.** As special events create the as-if randomness at the core of a quasi-experimental study, they can result in other special features that never threaten external validity.

In case of IV regression with heterogeneous causal effects, the two stage least square estimator (TSLS) is a (1) _____ estimator of the weighted average of the individual causal effects also known as the local average treatment effect (LATE).

In which of the following cases is the local average treatment effect not equal to the average treatment effect?

- ○ **A.** When the heterogeneity in the treatment effect and heterogeneity in the effect of the instrument are uncorrelated.
- ○ **B.** When the treatment effect is the same for all individuals.
- ○ **C.** When the instrument affects each individual equally.
- ○ **D.** When an individual's decision to receive treatment depends on the effectiveness of the treatment for that individual.

Suppose two researchers use two different instruments that are both valid (i.e., both are relevant and exogenous) to estimate the average causal effect in heterogeneous populations.

In large samples, the researchers would obtain (2) _____ TSLS estimators of the average causal effect.

(1) ○ inconsistent ○ consistent

(2) ○ different ○ the same

Answers A.

Just because an instrument is randomly determined or as-if randomly determined in a quasi-experiment does not necessarily mean it is exogenous.

, B.

Checking for systematic differences between the treatment and control groups to test for failure of randomization is not possible as some factors in the error term which could be related to the treatment level *X* are unobserved.

(1) consistent

D. When an individual's decision to receive treatment depends on the effectiveness of the treatment for that individual.

(2) different

ID: Concept Exercise 13.6.2