

אוניברסיטת תל אביב, ביה"ס לכלכלה ע"ש איתן ברגלס

מבוא לאקונומטריקה - תרגיל מסכם

סמסטר ב' תשפ"ב

נועה דה לה וגה

מטרת העבודה:

מטרת העבודה הינה ביצוע מחקר אמפירי תוך שימוש בכלים אקונומטריים שלמדתם בקורס: רגרסיה ליניארית, בדיקת השערות, משתני דמי, וכדומה.

הוראות כלליות:

מועד ההגשה: 25.7.2023 (לא יתקבלו עבודות המוגשות באיחור)

ההגשה תעשה בעותק אלקטרוני בלבד. את העותק האלקטרוני יש להגיש בתיבת ההגשה ב-moodle (פעם אחת בלבד לכל זוג).

ההגשה היא בזוגות בלבד. העבודה תהווה 20% מהציון הסופי בקורס.

על מנת לקבל ציון עובר בקורס יש לקבל ציון עובר גם במבחן וגם בעבודה זו.

עד עתה בתרגילים קיבלתם הוראות מפורטות, כאשר התבקשתם לאמוד מודלים ספציפיים תוך כדי שנאמר לכם מה לכלול במודל ומה לא, איזו השערה לבדוק וכו'. בחלקים רבים בעבודה זו אתם מתבקשים להחליט באופן עצמאי (תוך שימוש בכלים שקיבלתם בקורס) מהי רשימת המשתנים המסבירים שתכלול במודל, את הצורה הפונקציונאלית של המודל ואת ההשערות לבדיקה.

בנוסף:

- הגבילו את אורך העבודה ללא יותר מ-15 עמודים (ואפשר אפילו הרבה פחות), פונט - David 12 עם רווח של 1.5 (לא כולל נספחים / קבצי Rscript).
- הסעיפים מוספרו על מנת להקל על כתיבת וארגון העבודה. הציון יינתן על בסיס הערכה של איכות העבודה כמכלול.
- יש להגיש בנספח העבודה תדפיס של הקוד.
- כתיבה וניסוח ברור ופורמאלי מהווים מפתח חשוב להצלחה בעבודה. הימנעו מפרשנות עמומה ולא מדעית.
- במבחני השערות נסחו את השערות האפס באופן פורמאלי והציגו את חישוב הסטטיסטי לבדיקת ההשערה. במבחני t , F הציגו את הערך הקריטי שמעליו דוחים את השערת האפס. לביצוע המבחנים עליכן להסתמך על רמת מובהקות של 5%.
- תנו פרשנות מדויקת לאומדים והבליטו את התוצאות הרלוונטיות מכל פלט.
- מומלץ לקרוא את כל המסמך לפני שמתחילים לעבוד עם הנתונים.

בהצלחה!!!

העבודה מתבססת על המאמר הבא:

Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review*, 95(1), 414-424.

בשנת 1940 בשוודיה תוכננה רפורמה נרחבת במערכת החינוך שמטרתה הייתה לשפר את הגישה לרמות השכלה גבוהות יותר. במסגרת הרפורמה הוגדלו מספר שנות לימודי החובה מ 7/8 ל 9 שנים. כמו כן בוטל אופן שיבוץ התלמידים ללימודים אקדמיים ולא אקדמיים על בסיס ציונים לאחר כיתה ו' והוחלט על יישום תוכנית לימודים אחידה כלל ארצית. ליישום הרפורמה קדם ניסוי שבמהלכו הרפורמה יושמה באופן מדורג בין עיריות שונות. לרשות החוקרים עמדו נתוני סקר אודות אנשים משני שנותוני לידה (אנשים שנולדו ב 1948 ואנשים שנולדו ב 1953), כאשר בחלק מהעיריות נערכה הרפורמה במהלך שנים אלו (כלומר מי שנולד ב 1948 היה תחת מערכת החינוך הישנה ומי שנולד ב 1953 היה תחת מערכת החינוך החדשה) ואילו בחלק מהעיריות הרפורמה נערכה בשנים מאוחרות יותר (כלומר שילדי 1948 וילדי 1953 שניהם היו תחת מערכת החינוך הישנה).

לרשות החוקרים עמדו גם נתונים אודות עיריות שבהן הרפורמה נערכה קודם לכן (כלומר שילדי 1948 וילדי 1953 שניהם היו תחת מערכת החינוך החדשה). אולם מהנתונים שלכם השמטנו את הקבוצה הזו על מנת להקל עליכם.

שימו לב שבאופן כללי - על מנת לפשט את הדברים תתבקשו להעריך באופן שונה את שאלת המחקר ואין הכרח לקבל את אותן התוצאות שקיבלו כותבי המאמר.

*****לפני שאתם ממשיכים מכאן - יש לקרוא בעיון את המאמר עליו מבוססת העבודה*****
*****קריאת המאמר הכרחית על מנת לענות על השאלות שבהמשך העבודה*****

חלק א' – הצגת שאלת המחקר וחשיבותה הכלכלית

הציגו את שאלת המחקר במאמר.

1. הסבירו במילים שלכם, מה מטרת המאמר? מהי שאלת המחקר?
2. איך המחקרים מציעים לבדוק את שאלת המחקר? באיזו אסטרטגיה אמפירית הם משתמשים?
3. מהן לטענת החוקרים ההנחות עליהן האסטרטגיה האמפירית מבוססת?
4. האם הגרף שמוצג במאמר (FIGURE 1) מחליש או מחזק את ההנחות עליהן מתבססים החוקרים?
5. הסבירו בקצרה מהן התוצאות העיקריות של המאמר (אין צורך לציין את כל התוצאות, יש להתרכז בממצאים העיקריים).

הערה חשובה: ההסבר צריך להיות במילים שלכם – לא יתקבלו תרגומים של משפטים מהמאמר לעברית.

חלק ב' – ביצוע עבודה אמפירית

6. כל מאמר מתחיל עם סטטיסטיקה תיאורית של הנתונים. בהתבסס על הנתונים של פרטים שנולדו ב-1948 ועבור שנת 1985, צרו טבלה בה אתם מדווחים על הממוצעים של המשתנים העיקריים עבור פרטים המתגוררים בעיריות שבהן שיטת החינוך הוחלפה ובין פרטים שמתגוררים בעיריות ששיטת החינוך בהן לא הוחלפה בנפרד (2 עמודות). הוסיפו לטבלה עמודה נוספת (עמודה 3) אשר תציג את ההפרש בממוצעים בין שני סוגי העיריות וסטיית התקן של ההפרש + עמודה אחרונה (עמודה 4) המציגה את ה p -value של השערה הבודקת האם ההפרש בין הממוצעים מובהק. רמז: חשבו איך לקבל את התוצאות של עמודה 3 ו-4 באמצעות סדרה של רגרסיות ובצעו זאת.
7. האם התוצאות שקיבלתם מפתיעות אתכם נוכח העובדה שהרשויות שיישמו את שיטת החינוך החדשה לא נבחרו באופן אקראי? מדוע?
8. צרו טבלה כמו בסעיף 6 עבור פרטים שנולדו בשנת 1953. דונו בתוצאות שלה בקצרה כמו בסעיף 7.
9. כעת דונו בהשלכות של סעיפים 6-8 על האסטרטגיה האמפירית של הכותבים במאמר. האם אלו עדויות שמחזקות או מחלישות את האמינות שלה.
10. על בסיס הנתונים של הפרטים שנולדו ב-1948, נסחו מודל אקונומטרי המנסה לבחון את הקשר בין שנות הלימוד של הפרטים (years) וההכנסה שלהם בשנת 1985. כתבו את המודל באופן מפורש ואמרו אותו. אתם צריכים להחליט לבד איזה משתנים מסבירים להוסיף ואת הצורה הפונקציונאלית (לוג, רמה, ריבועי וכו'). הצדיקו את בחירתכם. הציגו את תוצאות האמידה בטבלה ופרשו את משמעות המקדמים תוך התייחסות למובהקות שלהם, משמעותם הכלכלית ולמובהקות של המודל בכללותו.
11. בהתבסס על המודל שניסחתם בסעיף 10, האם האומד לשנות השכלה צפוי לתאר את הקשר הסיבתי שבין שנות השכלה לשכר? הסבירו.
12. עתה אמרו את המודל בסעיף 10 (עם אותו מדגם, פרטים שנולדו בשנת 1948 והכנסתם ב-1985) רק שהפעם השתמשו במשתני דמי לרמות ההשכלה השונות (צרו קטגוריות לפי ראיות עיניים - אפשר להשתמש לצורך יצירת הקטגוריות במשתנה level). דונו באומדנים שקיבלתם למשתני ההשכלה השונים.
13. מהי ההנחה המובלעת בניסוח המודל בסעיף 10? מהי ההנחה במודל בסעיף 12? איזה מודל לדעתכם מתאר טוב יותר את המתאם בין השכלה לבין שכר (המודל בסעיף 10 או המודל בסעיף 12)?

הן אלו
הנחות
מחיל
והאם

חלק ג' – כעת נתרכז בשאלה: האם הרפורמה במערכת החינוך השוודית השפיעה על ההשכלה והשכר של הפרטים?

****שימו לב שעבור סעיפים 14-21, יש להשתמש בתצפית אחת לכל פרט ולכן השתמשו בסעיפים אלו רק**

בנתונים עבור שנת 85

14. חשבו את הממוצע של שנות הלימוד בנפרד עבור פרטים שנולדו ב-1948 ו-1953, לפי האם באיזור שלהם הרפורמה לא חלה על אף שנתון (משתנה treat_muni שווה ל-0) ולפי האם באיזור שלהם הרפורמה חלה

האם זהו אפקט טיפה?
אם כן, איך זה ייראה?

- על שנתון 1953 ולא על 1948 (משתנה treat_muni שווה ל-1). צרו בעזרת 4 ממוצעים אלו טבלה של הפרש הפרשים, ודווחו על ההשפעה הסיבתית הנאמדת של הרפורמה על שנות לימוד לפי חישובכם.
15. נסחו מודל אקונומטרי של הפרש הפרשים (ללא משתנים מפקחים נוספים) על מנת לבדוק את השינוי שחל **בסה"כ שנות הלימוד של הפרטים** כתוצאה מהרפורמה. כתבו את משוואת המודל במפורש. דווחו על התוצאות של האמידה בטבלה והסבירו את המשמעות של כל אחד מהאומדים (במונחים סטטיסטיים ובמונחים כלכליים) תוך התייחסות למובהקות שלהם ולגודלם.
16. מהי ההנחה המרכזית של המודל שאמדתם לצורך קבלה של השפעה סיבתית? (הסבירו את ההנחה באופן כללי ובהקשר של המודל הספציפי שאמדתם). מה דעתכם על סבירות ההנחה? (הסבירו).
17. בדקו האם ההנחה של שונות שווה סבירה במודל שאמדתם בסעיף 15 בעזרת המבחן המתאים. במידה וההנחה אינה סבירה, הסבירו מהן ההשלכות מהתעלמות מבעיה זו, תקנו ודווחו על התוצאות.
18. הוסיפו למודל הפרש הפרשים מסעיף 15 משתנים מפקחים על פי ההיגיון שלכם (ניתן גם להיעזר בבחירה של המחבר במאמר). כתבו את המודל ואמדו את המשוואה מחדש (תוך תיקון השונות במידה ויהיה צורך). הצדיקו את בחירתכם של המשתנים המסבירים על בסיס הגיון כלכלי.
19. בדקו האם המשתנים המסבירים שהוספתם בסעיף 18 מובהקים (ביחד) – כלומר, האם היה מוצדק להוסיף אותם. הציגו את ההשערה, את סטטיסטי המבחן, את תוצאת המבחן ומסקנתכם על פי התוצאות.
20. דווחו על המקדם העיקרי של המודל שקיבלתם בסעיף 18 והסבירו את משמעותו (תוך התייחסות למובהקות סטטיסטית, גודל ומשמעותו הכלכלית). האם הוא שונה מהמקדם שקיבלתם בסעיף 15? הסבירו את ההבדל.
21. מהי ההנחה המרכזית של המודל שאמדתם בסעיף 18 לצורך קבלה של השפעה סיבתית? (הסבירו את ההנחה באופן כללי ובהקשר של המודל שאמדתם). מה דעתכם על סבירות ההנחה? (הסבירו). במה היא שונה מההנחה של המודל מסעיף 15, שנידונה בסעיף 16?
22. כעת אמדו מודל דומה לזה שניסחתם בסעיף 18, אבל הפעם על משתנה התוצאה לוג הכנסה השנתית של הפרטים. חישובו והסבירו **אילו משתנים חשובים שתפקחו עליהם עבור אמידה זו**. כתבו את המודל באופן מפורש כולל כל המשתנים המפקחים. דווחו על המקדם העיקרי של המודל והסבירו את משמעותו.
23. כעת אמדו את המודל מסעיף 22 בנפרד עבור גברים ונשים. לאחר מכן, בצעו חלוקה נוספת ואמדו את שני המודלים (גברים ונשים) בנפרד עבור פרטים שיש להם אב בעל השכלה גבוהה (למד יותר שנים מחינוך חובה) ואב בעל השכלה נמוכה (4 מודלים בסה"כ). סדרו את התוצאות שקיבלתם (6 מודלים בסה"כ) בטבלה בדומה לטבלה 2 במאמר - אין צורך לבצע את החלוקה לפי יכולות הפרטים. מה דעתכם על התוצאות שקיבלתם? כיצד תוכלו להסביר?
24. אחד האתגרים המרכזיים של המאמר, הוא הצדקת ההנחה לקבלת אפקט סיבתי. אילו דברים נוספים הייתם מציגים על מנת להצדיק את ההנחה הזאת בהקשר של מודל הנאמד במאמר? הניחו שאין לכם שום מגבלה לאיסוף נתונים. (אין צורך בהצגת הממצאים, רק תיאור הממצאים / בדיקות שהייתם מוסיפים).

תצורה
זכא

Code Book for Data Set

id - individual identifier

lnearn - log annual nominal labor earnings.

year - year (1985-1996) of earnings.

d53 - 1 for born in 1953, 0 otherwise.

treat_muni – 1 if municipality assigned cohort 1953 to the reform (and didn't 1948), 0 otherwise

female - 1 for females, 0 for males.

hiab3 - 1 for individuals with above average scores on ability tests, 0 otherwise.

hife - 1 for individuals with father's with more than compulsory education, 0 otherwise.

level - education levels 1-7. 1 is lowest.

- 1 - Pre-reform compulsory school
- 2- Post-reform comprehensive (compulsory school) or pre-reform junior secondary school,
- 3- Upper secondary school ≤ 2 years,
- 4- Upper secondary school ≥ 3 years,
- 5- Post upper secondary school ≤ 2 years,
- 6- Post upper secondary school ≥ 3 years (University/College),
- 7- Ph.D. or licentiate degree at a University).

yearscho - number of years of schooling

fk - code for municipality of schooling

q1 –

- q1=1 if the municipality of birth was assigned to the reform
- q1=2 if the municipality of birth was not assigned to the reform
- q1=3 if the municipality of birth was ambiguously assigned to the reform.