Homework 4

Introduction ot Econometrics,Fall 2017

Zhaopeng Qu

12/13/2017

1 Due Date:

- ***Due to Tues. Dec. 26 11:00 p.m.***

2 Learning Objectives

- Use *Stata* or *R* to understand and practice following Econometric Methods
  - **LPM,Logit and Probit Model**
  - **Oaxaca-Blinder Decomposition**

3 Main Content

3.1 Review: LPM,Probit and Logit Models(Chapter 11)

- **Analysis Exercises: 11.1-11.5** (pp450-451 in *SW textbook*)
- **Empirical Exercises: E11.1-E11.2** (P452-453 in *SW textbook*)

3.2 Oaxaca-Blinder Decomposition

- In this exercise, you are asked to produce some results about Rural-urban migrants v.s Urban natives in China using a cleaned subset of *Rural to Urban Migrant Suvey in China(RUMiC)*.Please download from here **CHIP2002_2007**

1. Generate basic summary statistics such as mean, standard deviation, and total number of observations for all variables(including **age,gend,enth,masta,edu,hwage and exper**) in the data sets by migrate status(thus one for migrants and the other for urban natives) in two years(2002 and 2007) respectively.
2. Make comparisons about these variables between migrant and urban natives across two years and use a simple t-test to present if there is a significant mean difference between these variables in two years(2002 and 2007) respectively. Try combine the result with question 1 in just one table whose structure is like the following table(**Table 1**)

Variables	urban	migrant	difference
hourly earnings	mean (SD)	mean (SD)	difference [s.e]

.

Variables	urban	migrant	difference
the Number of Obs.			

3. The classical Mincer wage equation is

$$lny_i = \beta_0 + \beta_1 edu + \beta_2 exper + \beta + 3exper^2 + \gamma X + u_i$$

,

- where X include marital status and minority variable. We put some dummy variables to run a series of OLS regression to measure the log wage differentials as following one by one

  1) a dummy variable of migrating status(migrants/urban residents) in 2002 and 2007 respectively.
  2) a dummy variable of gender (male/female) in 2002 and 2007 repectively
  3) a dummy variable of selecting year(2002/2007).
  4) add an interaction term of migrant and gender based on 1) in 2002 and 2007 respectively.
  5) add intercation terms of migrant+gender,year+gender,year+migrants and migrant+year+gender based on 3)

- Please put these regressions into one table(**Table 2**), report the amounts and the significance of these estimate coefficients of dummies and interactions.And try to present the economic implications of these results respectively.

4. Use Stata/R command to perform an Oaxaca-Blinder decomposition(after being sure that the oaxaca package has been installed on your computer) males and females in 2002 and 2007 wage gaps and the dynamics across 2002 and 2007.

   1) Please use in turn the urban coefficients, the migrants coefficients and two weighted(0.5 and sample-adjusted) average of the two in 2002answer the following questions: Report the amount and significance of explained effects and unexplained effects, and calculate the shares of these effects to total wage gaps. Put all results in one table(**Table 3**) and presents which one dominate more in 2002.

   - Based on **Table 3**,answer the following questions:a) how to explain the fact? b) Is there some differences among these results using different weights? Please explain why?

   2) Please use in turn the Oaxaca and Ransom(1994) or Jann(2008) weighted average of male and female in 2002 and 2007 respectively.Report the amount and significance of explained effect and unexplained effect for three specific variables such as education,age and experience. Calculate the shares of these effects to total explained/unexplained effects and to total wage gaps,receptively. Put all results in one table(**Table 4**)

   - Based on **Table 4** answer the following questions: a)which one dominate more for three variables in 2002 and 2007,respectively? b)Please explain these results one by one in 2002 and 2007. c) does anything change during 2002-2007? Explain what is going on.

   3) Please use OB decomposition with Jan(2008)'s weight to distengle unexplained and explained effect to the wage change across 2002 and 2007. Report the amount and significance of explained effect and unexplained effect for total *AND* three specific variables such as education,age and experience. Calculate the shares of these effects to total explained/unexplained effects and to total wage changes during 2002-2007,repectively. Put all results in one table(**Table 5**)

   - Based on **Table 5** answer the following questions:a) explain these results in detail. 2) make a comparsion with the result of the last subquestion in question 3(thus "And does anything change during 2002-2007? Explain what is going on?) and tell what is the difference of implications between these two decomposition results.