Labor Economics, Spring 2021

Problem Set 3
Due Friday, June 16 in class

1. [15 points] Write a summary on the following articles from the New York Times:

   https://www.nytimes.com/2006/07/09/magazine/the-immigration-equation.html or the

   following episodes from the Freakonomics Radio:

   https://freakonomics.com/podcast/minimum-wage/

   https://freakonomics.com/podcast/the-true-story-of-the-gender-pay-gap-a-new-freakonomics-radio-podcast/.

   Your summary should follow the following requirements:

   - Double-spaced, 1 inch margins, a standard 12 pt font
   - Length no longer 500 words

2. [15 points] The following questions are about measuring the returns to another year of schooling.

   a. Short answer: Why can't we compare the average wage of high school and college graduates to arrive at a measure of the returns to schooling?

      High school and college graduates may be different in many ways, including family background, neighborhoods, innate ability, etc. Therefore, their earning difference is a combination of the return to schooling and their differences in characteristics.

   b. Short answer: How might the use of twins be helpful in measuring the returns to schooling? Are there any potential problems?

      Twins are the same genetically, so they may have similar innate ability. However, they may be different in their current ability, leading to different years of education. Therefore, comparing twins is helpful to control innate ability, but it may not be able to control for their current ability.

   c. Ashenfelter proposed to estimate the causal effect of schooling on income based on the Vietnam draft lottery—the system where random numbers were assigned to birthdates and then the lower numbered days were drafted first. In short, it was observed that individuals with lower draft lottery numbers went to school longer to avoid being drafted. The following table shows the average education and annual salary of two groups of men: those born on September 7, 1950, and those born on September 8, 1950. (a) Based on

Ashenfelter's methodology, what is an estimate of the rate of return to a year of schooling? (b) What is the key assumption of Ashenfelter's method?

Birth date	Lottery Number	Average years of education	Average salary in 1990
September 7, 1950	1	12.3	$42,000
September 8, 1950	307	11.8	$40,000

The estimated return to a year of schooling is equal to ($42,000-$40,000)/(12.3-11.8)=$4,000. The key assumption here is that individuals who were born on September 7, 1950 are similar to those born on September 8, 1950. For example, this method will underestimate the return to education if individuals assigned a lower number is also more likely to be drafted, which might lower their future earnings.

3. [15 points] This exercise will use the 2014 Manpower Utilization Survey to estimate Mincer style earning function. Here are the links for 2014 MUS data:

   https://www.dropbox.com/s/ldvi3x2utm9v3v6/2014%20Manpower%20Utilization%20Survey.dta?dl=0

   and the codebook:

   https://www.dropbox.com/s/g3a0wrue12nx7z3/MUS%202014_Code%20book.pdf?dl=0

   Limit the sample to working individuals (*Hint: a8==1*) who age at least 25 when interviewed.

   a. What is the fraction of individuals who have no college degree? What is the fraction of individuals who have at least a college degree? (*Hint: use tab command*)

      As seen in the tables below, about 59% of individuals do not have a college degree, while 41% of individuals have a college degree.

```
. keep if a8==1 & a3>=25
(29,504 observations deleted)

. gen col=a5_2>=7

. mean col

Mean estimation                   Number of obs    =     26,763
```

	Mean	Std. Err.	[95% Conf.	Interval]
col	.4098569	.0030063	.4039643	.4157494

   b. Report the average earnings for high-school graduates and the average earnings for individuals who have at least a college degree. Are they significantly different from each other at 5% level? (*Hint: use ttest command*)

According to the table below, the earning difference is significantly different from zero at 5% level.

```
. ttest b1_a if a5_2>=5, by(col)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	9,162	31832.14	301.7499	28883	31240.64	32423.63
1	10,969	43669.86	309.3826	32402.57	43063.41	44276.31
combined	20,131	38282.29	221.3645	31408.02	37848.39	38716.18
diff		-11837.72	436.6339		-12693.56	-10981.89

```
    diff = mean(0) - mean(1)                                  t = -27.1113
Ho: diff = 0                     degrees of freedom =    20129

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

c. Estimate a Mincer style earnings function by including the following explanatory variables: education, age, age squared, and gender. What is your estimated return to education? Do you think there is any causal interpretation for your estimate?

The estimate below suggests that one year of additional education is associated with 6% increase in earnings. The estimated return of education is unlikely to be causal because the model makes no effort to control for ability.

```
. replace school=0 if a5_2<=2
(244 real changes made)

. replace school=6 if a5_2==3
(2,300 real changes made)

. replace school=9 if a5_2==4
(4,088 real changes made)

. replace school=12 if a5_2==5
(2,855 real changes made)

. replace school=12 if a5_2==6
(6,307 real changes made)

. replace school=16 if a5_2==7
(3,424 real changes made)

. replace school=16 if a5_2==8
(5,954 real changes made)

. replace school=18 if a5_2==9
(1,432 real changes made)

. replace school=23 if a5_2==10
(159 real changes made)


. gen asq=a3*a3

. gen lnw=log(b1_a)
(1,809 missing values generated)


. reg lnw school a3 asq a2, r

Linear regression                               Number of obs     =     24,954
                                                F(4, 24949)       =    1126.22
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1839
                                                Root MSE          =      .5385

------------------------------------------------------------------------------
             |               Robust
         lnw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      school |   .0660353   .0012053    54.79   0.000     .0636729    .0683977
          a3 |   .0510176   .0026366    19.35   0.000     .0458497    .0561856
         asq |  -.0004854   .0000307   -15.79   0.000    -.0005457   -.0004252
          a2 |  -.2218229   .0064709   -34.28   0.000    -.2345062   -.2091397
       _cons |   8.618532   .0570495   151.07   0.000     8.506712    8.730353
------------------------------------------------------------------------------
```