8686- Problem Set #5- Ashenfelter and Krueger's Twins Data

Twins.dta contains a sample from the Twins data used in Ashenfelter and Krueger *AER* 1994. The sample contains all white identical twins in the original data set. All observations contain a wage report for the last job worked (twins were dropped from the sample if one of them had not worked in the past two years). Variable descriptions and means are printed on the following page. Once you read in the data, take a look at it in order to make sure you understand the data structure. You might find it helpful to list the four education variables: *educ11, educ12, educ21*, and *educ22* in order to make sure you understand their definitions. You will almost certainly need to refer to the *AER* article and class discussion of that article to complete the Problem Set

1) Create the variable *wage* (unlogged) from the *lnwage* variable.
Regress *wage* on own report of own education (*educ11*)
Regress *lnwage* on the same variable.
Calculate the residuals from each regression.
In separate graphs, plot each set of residuals against *educ11*.
Based on the regression results and what you see in the plots, which do you prefer as the dependent variable: *wage* or *lnwage*? Why?

For what follows, continue to use *lnwage* as the dependent variable and *educ11* as the independent variable of interest.

2)Run a regression adding the *age* variable and then another adding both *age* and *agesq*.
What does this show about the relationship between age and wage?
How does adding the age controls affect the education coefficient and what does that tell you about the relationship between age and education? Check this directly.

3) Run a regression that includes *age*, *agesq*, *marr* and *female*.
Run a second regression that adds the interaction term *marrfe=marr*female*.
Interpret the results.
Is the relationship between wage and marriage the same for men and women?
Can you explain this result?

4) Estimate a standard returns to education regression with controls for age, gender, marital status and union coverage.
Interpret the magnitude of the returns to education.

5) Estimate a differences specification, using the difference in twins' wages as the dependent variable and the difference in education levels (using each sibling's own report of their own education) as the independent variable. Do not include any controls.

6)Add controls to the specification used in #5.
Recognize that now you are controlling for the differences in the variables, so some control variables will drop out (which ones?).
How does the estimate of the returns to education compare to the estimate obtained in #4?
Is it surprising that the coefficient changed in this direction?

7) Estimate a regression that corrects for both family fixed-effects and classical measurement error. Be clear about the specification you are using and why this specification should correct bias due to classical measurement error. How does this change your estimate of the return to education?

8) Estimate a regression that corrects for both family fixed-effects and non-classical measurement error that might exist if reports by the same twin tend to have correlated measurement errors (in other words, if a person tends to brag about both themselves and their sibling). Be specific about the specification you are using and why this specification should correct bias due to this form of measurement error. How does this change your estimate of the return to education?

```
Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+--------------------------------------------------------
  pairid |     284    240.7324   145.4565          2        499
   order |     284         1.5   .5008826          1          2
     age |     284     36.4422   10.25805   18.78166   64.13963
    ucov |     284    .2288732    .420849          0          1
  female |     284    .5633803   .4968422          0          1
  educ11 |     284    14.14437   2.119305         10         20
  educ12 |     284    14.05986    2.11672         10         20
    marr |     284     .443662   .4976929          0          1
  marrfe |     284    .2711268   .4453261          0          1
  lnwage |     284    2.355952   .6085004   .5108258    4.60517
   agesq |     284    1432.891   826.1881   352.7508   4113.893
  dlwage |     142    .0497417   .5607356  -1.446919   2.109747
 lnwage2 |     142    2.331081   .6260922   .5108258    4.60517
  educ22 |     142    14.16901   2.110332         10         20
  educ21 |     142    14.09859    2.13485         10         20
  deduc1 |     142   -.0492958   1.823127         -5          7
  deduc2 |     142   -.0774648   1.894651         -6          7
   dmarr |     142   -.0422535   .5035088         -1          1
 dmarrfe |     142   -.0211268   .4205419         -1          1
   ducov |     142    .0070423   .4837278         -1          1

Contains data from twins2.dta
  obs:           284                          Twinsburg white identicals
 vars:            20                          22 Mar 1995 16:15
 size:        24,708 (97.5% of memory free)
------------------------------------------------------------------------------
   1. pairid      int    %8.0g                 id number of pair
   2. order       byte   %8.0g                 order in pair (arbitrary)
   3. age         float  %9.0g                 age in years
   4. ucov        byte   %8.0g                 =1 if union coverage
   5. female      byte   %8.0g                 =1 if female
   6. educ11      byte   %8.0g                 own report of own education
   7. educ12      byte   %8.0g                 sibs report of own education
   8. marr        byte   %8.0g                 =1 if married
   9. marrfe      byte   %8.0g                 =1 if married female
  10. lnwage      float  %9.0g                 log wage
  11. agesq       float  %9.0g                 age squared
  12. dlwage      float  %9.0g                 lwage1-lwage2
  13. lnwage2     float  %9.0g                 log wage of twin 2
  14. educ22      byte   %8.0g                 sibs report of sibs education
  15. educ21      byte   %8.0g                 own report of sibs education
  16. deduc1      byte   %8.0g                 educ11-educ22
  17. deduc2      byte   %8.0g                 educ12-educ21
  18. dmarr       byte   %8.0g                 marr1-marr2
  19. dmarrfe     byte   %8.0g                 marrfe1-marrfe2
  20. ducov       byte   %8.0g                 ucov1-ucov2
```