**Lab 3 Stata**

This lab estimates the wage gaps between black and whites and Hispanics and whites in the US. The lab uses longitudinal cross-sectional data for a random sample of 6,808 individuals. The wage gaps are estimated in an standard Mincer framework. We analyze whether the results change when including a cognitive test score to the earnings equation. Secondly, to analyze if there are differences in the returns to education and returns to experience separate earnings equations will be estimated for the different groups. Finally, with a Oaxaca decomposition we divide the wage gaps into an explained and an unexplained part.

The laboratory assignment should be in the form of a short paper. That is, the results should be presented in tables and explained correctly. (If you hand in only the estimation results you will not pass).

*i)* Generate dummy variables for black and Hispanic, male, the different regions and marital status. Create variables for potential experience, potential experience2 and logarithmic wage.

Report summary statistics (mean and standard deviation) for blacks, Hispanics and whites.

*ii) The wage gap*

Estimate, separately for men and women, the baseline (without controls) wage gap between blacks and whites and between Hispanics and whites. Then estimate the gaps when controlling for:

1.Years of schooling, potential experience, and potential experience2

1. Years of schooling, potential experience, potential experience2, marital status, urban/rural, and region
2. Test score (only)
3. Years of schooling, potential experience, potential experience2, marital status, urban/rural, region, and test score

Why do the wage gaps change when more control variables are added? Which specification do you prefer and why?

iii) *Differences in coefficients:*

Estimate specification 2 separately for whites, blacks and Hispanics (and separately for men and women). Why do the coefficients change compared to i)?

Also, estimate specification 2 without potential experience2. Compare the differences in coefficients. Why does the potential experience estimate change when adding potential experience2 to the specifications?

iv) *Oaxaca-decomposition*Use the Oaxaca technique to decompose the wage gap between male blacks and male whites. Estimate the simplified wage equations for blacks and whites:

*Lnwi* =α*w* +β*wSi* +ε

*Lnwi* =α*b* +β*bSi* +ε

and calculate the explained difference and the unexplained difference. Instead of reporting the regression result in a table, draw a graph of your findings (like the textbook does on p. 385). We do not want you putting a lot of effort into making a nice graph, it is sufficient to draw a not-correctly-scaled graph with pen and paper, take a photo with your phone, and attach it to the report. Make a table where you summaries the raw wage gap, the explained difference and the unexplained difference. Explain and discuss the result!

**Useful Stata commands for this Lab**

*reg depvar. indep. variables if subgroup==X* OLS for subgroup X

or (if the subgroup is based both on the variable subgroup1 and the variable subgroup2) :

*reg dep.variable indep.variables if subgroup1==X & subgroup2==Y*

*ssc install Oaxaca* Might need to install the Oaxaca decomp.

*oaxaca depvar. indep. variables, by(decomp. var) noisily* Oaxaca decomp.

When running the above oaxaca command the output will be a bit different from what you  
expect from the textbook (p. 384). It relates as the following:

difference: is the raw wage gap

coefficients: is the “differential due to differences in skills”

endowment + interaction: is “differential due to discrimination”

We add the noisily option in the command so you can see the two models that underlies the decomposition. See B. Jann’s (2008) “A Stata implementation of the Blinder-Oaxaca decomposition” for more information about using the Oaxaca decomposition in Stata. https://ideas.repec.org/p/ets/wpaper/5.html

**Variables:**

**RACE**

1 HISPANIC  
2 BLACK  
3 NON-BLACK, NON-HISPANIC

**SEX**

1 MALE  
2 FEMALE

**REGION OF CURRENT RESIDENCE**

1 NORTHEAST  
2 NORTH CENTRAL

3 SOUTH  
4 WEST

**CURRENT RESIDENCE URBAN/RURAL?**

0 RURAL

1 URBAN

**MARITAL STATUS**

1 NEVER MARRIED  
2 MARRIED, SPOUSE PRESENT

3 OTHER

**HIGHEST GRADE COMPLETED AS OF MAY 1 SURVEY YEAR**

1 1ST GRADE  
2 2ND GRADE  
3 3RD GRADE  
4 4TH GRADE  
5 5TH GRADE  
6 6TH GRADE  
7 7TH GRADE  
8 8TH GRADE  
9 9TH GRADE  
10 10TH GRADE  
11 11TH GRADE  
12 12TH GRADE  
13 1ST YEAR COLLEGE

14 2ND YEAR COLLEGE

15 3RD YEAR COLLEGE

16 4TH YEAR COLLEGE

17 5TH YEAR COLLEGE

18 6TH YEAR COLLEGE

19 7TH YEAR COLLEGE  
20 8TH YEAR COLLEGE OR MORE

95 UNGRADED