

Problem Set V

MIT (14.32)
Spring 2007

A. From Wooldridge: 8.1, 8.2, 8.5, C8.5, 10.1, C10.5, 12.1, C12.10

B. Additional problems:

This problem set asks you to replicate Krueger's 1993 QJE paper, "How Computers Have Changed the Wage Structure." The paper can be downloaded from JSTOR. CPS extracts similar to those Krueger used are posted on the course website in SAS format. Many variables are constructed already, but you should note that Krueger uses sample restrictions (e.g. age, work status, maximum/minimum wages allowed in sample, etc.) that you will need to replicate (see Appendix A of his paper for details, but note that the \$999 vs. \$1,923 top-coding issue is already taken care of).

You should attempt to reproduce the results as closely as possible. Note, however, that you should not expect to get *exactly* the same results/sample size as Krueger – there are simply too many choices to make on an empirical research project! But all coefficients should be of the same sign, significance, and close in magnitude to Krueger's. For example, a coefficient of 0.090 with a standard error of (0.02) would be considered close to a coefficient of 0.098. A coefficient of 0.050 with a standard error of (0.02) would not be considered close to a coefficient of 0.098.

Include SAS logs and listings when submitting your problem set. Also, organize your results into three tables with the same layout as the three Krueger tables, and list Krueger's results next to your own (i.e. one column of your results, and then the corresponding Krueger column next to it, and then the next column of your results, and then next Krueger column, and so on).

Specific tasks

Replication

1. Reproduce all Table I results, with the exception of the occupation means (the definition of these variables is not clear). Also note that Krueger's part-time variable does not match ours (but everything else should match).
2. Reproduce all Table II results. Everything should match closely except for the "other race" variable.
3. Reproduce all Table III results. Everything should match reasonably closely.

Explore variations

4. Run the regression in Table II, Column (6) without region dummies. Do the results change much? What do the numbers in Table I suggest the consequences of this are likely to be?
5. Is serial correlation an issue in these models? Why or why not? What about heteroscedasticity? For the results in Table II, columns 1 and 4, report regular and heteroscedasticity-consistent standard errors.
6. Building upon Krueger's specification in Column 6, Table II, let the computer-use effect vary by union status. Test whether it does. Let both the computer-use and schooling effects vary by union status. Test these interactions for statistical significance and briefly discuss the results.
7. We have seen in class that blacks may have a higher return to education than whites. Construct and run a regression model that allows both the computer use and schooling effects to vary in the two nonwhite race groups (i.e., black vs. white, other vs. white). Construct F-tests for these differences, one for each race group and jointly. How might you have expected the computer-use effect to vary by race?