BOSTON COLLEGE
Department of Economics
EC771: Econometrics
Spring 2004
Prof. Baum, Ms. Uysal

PROBLEM SET 5: DUE TUESDAY 6 APRIL 2004 AT CLASSTIME

All references to Greene, 5th ed., 2003.

1. Problem 7.1, using

```
use http://fmwww.bc.edu/ec-p/data/greene2000/TBL8-1.dta
```

2. Problem 7.2

3. With the Stata data set

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/wage2.dta
```

a. Estimate the regression of log wage on education, experience, tenure, and dummies for married, black, south, and urban. Ceteris paribus, what is the approximate difference in monthly salary between blacks and nonblacks? Is this difference statistically significant?

b. Extend the model to allow the return on education to depend on race, and test that hypothesis.

c. Starting with the original model, allow wages to differ across four groups: married/black, married/nonblack, nonmarried/black and nonmarried/nonblack, taking account of the interactions between those characteristics. What is the estimated wage differential between married blacks and married nonblacks? (Hint: `lincom` is useful in reporting these results).

(over)

4. With the Stata data set

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/gpa2.dta
```

a. Consider the regression of $colgpa$ on $hsize$, $hsize^2$, $hsperc$, sat, $female$ and $athlete$ where these variables are, respectively, college grade point average, size of high school graduating class, academic percentile in high school class, SAT score, and dummies for female and athlete. What are your expectations about the coefficients in this equation?

b. What is the estimated GPA differential between athletes and nonathletes? Is it statistically significant?

c. Drop sat from the model and reestimate the equation. Does this change the estimated effect of being an athlete?

d. In the original model, allow the effect of being an athlete to differ by gender, and test the hypothesis that there is no difference in estimated GPA between female athletes and female nonathletes.

e. Does the effect of sat significantly differ by gender?

5. With the Stata data set

```
use http://fmwww.bc.edu/ec-p/data/wooldridge2k/401ksubs.dta
```

a. Test the hypothesis that average $nettfa$ does not differ by 401k eligibility status. What is the dollar amount of the estimated difference?

b. Estimate the regression of $nettfa$ on inc, inc^2, age, age^2, $male$ and $e401k$. Are the quadratic terms justified? What is the estimated dollar effect of 401k eligibility?

c. Add the interaction terms $e401k(age - 41)$ and $e401k(age - 41)^2$ to the model. Given that the average age is about 41, the effect of $e401k$ is the estimated effect at the average age. Are these interaction terms significant? Interpret how they alter the model. Do the estimated effects of 401k eligibility at age 41 differ much between this model and the model of part b?

d. Drop the interaction terms from the model, but define five family size dummies, $fsize1$–$fsize5$ where $fsize5 = 1$ if $fsize \geq 5$. Include these dummies in the model (choosing a base group). Are they jointly significant? How do you interpret the resulting regression?

e. Consider the regression of part b in its fully interacted ("Chow test") form, in which both intercepts and slopes are allowed to differ over the five family size categories. (Hint: you may want to use `xi`). Test the hypothesis that the regression equation is stable over family size categories.