# BONUS PROBLEM SET

**Professor Hungerman**
**Fall 2009**

1. You have a datset of 1,000 women over the age of 30, and you want to run a regression where the dependent variable is the number of children each woman has. You want to see if this depends on a woman's education, age, and a dummy that equals 1 if a woman is married:

$$children = \alpha_0 + \alpha_1 ed + \alpha_3 age + \alpha_2 marital + u$$

Furthermore, you have 4 racial groups of women in your data. Using dummy variables and interactions of dummy variables, you run a regression that allows for different intercepts and slopes for each racial group (you also allow the affect of marriage to vary for each racial group).

You then want to test whether the regression lines estimated for each racial group are different, or whether the data can be explained by a regression that restricts intercepts and slopes to be the same for all racial groups. You recall that a Chow test is suitable for this. What are the relevant numerator and denominator degrees of freedom for a Chow test in this case?

*Call the racial groups A, B, C, and D. The unrestricted equation is:*

$$\begin{aligned} children = {} & \beta_0 + \beta_1 ed + \beta_2 age + \beta_3 marital \\ & + \beta_4 A + \beta_5 ed*A + \beta_6 age*A + \beta_7 marital*A \\ & + \beta_8 B + \beta_9 ed*B + \beta_{10} age*B + \beta_{11} marital*B \\ & + \beta_{12} C + \beta_{13} ed*C + \beta_{14} age*C + \beta_{15} marital*C + u \end{aligned}$$

*where A is a dummy that equals one for race A, and similarly for B and C. The omitted group is D.*

*The restricted equation is* $children = \alpha_0 + \alpha_1 ed + \alpha_3 age + \alpha_2 marital + u$

*For each of your four racial groups you have four coefficients to estimate: an intercept and three slope coefficients. So you have 16 coefficients total (the betas are numbered up to 15, but we start at zero, so there are really 16 there). The degrees of freedom in your unrestricted equation is 1000-16 = 984. This is the denominator degrees of freedom.*

*Your restricted equation imposes the same intercept and slope for each group. Thus it has four coefficients to estimate. This equation restricts 12 coefficients to be zero (since we started with 16 and now have 4). So the number of restrictions = numerator degrees of freed = 12.*

2.  Consider a variable x that takes on the value 3 with probability ($n$ – 1) / $n$, and takes on zero otherwise.  What is the probability limit of x as $n$ goes to infinity? For a *given* $n$, what is the expected value of x?

*The probability limit is 3.  For a given* n, *the expected value is 3(n-1)/n .*

3. True/False: the assumption of homoskedasticity is needed for OLS estimates to be consistent.

*True.  Although we will discuss this more after the test.*

4. True/False: if an estimator is unbiased, then it is consistent too.

*False. It is possible for an estimator to be unbiased but inconsistent.*

5.  True/False: if our unobservable term $u$ is not normally distributed, then for a given set of $x$ variables it follows that the dependent variable $y$ will not be normally distributed either.

*This is true and was discussed in class (starting before the first midterm). We needed the assumption of normality to construct t-statistics and F-statistics but we will see next week that we can relax the assumption of normality as n grows large.*