

## Homework 10

### MRM Qualitative Regressors, Part 1 (15 points)

Instruction:

- This HW must be done in Rmarkdown!
- Please submit both the .rmd and the Microsoft word files. (Do not submit a PDF or any other image files as the TAs are going to give you feedback in your word document)
- Name your files as: HW10\_groupnumber\_name
- All the HW assignments are individual work. However, I highly encourage you to discuss it with your group members.
- The answer key will be uploaded on Canvas a couple of days after the due date.

**Question 1** Given the following:

$$\widehat{\log(bwght)} = 4.66_{0.22} - 0.004_{0.0009}cigs + 0.0093_{0.0059}\log(faminc) + 0.016_{0.006}parity + 0.027_{0.01}male \\ + 0.055_{0.013}white \\ n = 1,388, \quad R^2 = 0.0472$$

and:

$$\widehat{\log(bwght)} = 4.65_{0.38} - 0.0052_{0.0010}cigs + 0.0110_{0.0085}\log(faminc) + 0.017_{0.006}parity + 0.034_{0.011}male \\ + 0.045_{0.015}white - 0.003_{0.003}motheduc + 0.0032_{0.0026}fatheduc \\ n = 1,191 \quad R^2 = 0.0493$$

The variables are defined as:

*bwght* = birth weight, in lbs.

*cigs* = average number of cigarettes the mother smoked per day during pregnancy.

*parity* = the birth order of this child.

*faminc* = annual family income.

*motheduc* = years of schooling for the mother.

*fatheduc* = years of schooling for the father.

*male* = Whether the child is male.

*white* = Whether the child is white or not.

1. In the first equation, interpret the coefficient on the variable *cigs*. In particular, what is the effect on birth weight from smoking 10 more cigarettes per day?
2. How much more is a white child predicted to weigh than a nonwhite child, holding the other factors in the first equation fixed? Is the difference statistically significant?
3. Comment on the estimated effect and statistical significance of *motheduc*.
4. From the given information, why are you unable to compute the F statistic for joint significance of *motheduc* and *fatheduc*? What would you have to do to compute the F statistic? (Hint, look at the number of observations in each model. Restricted vs Unrestricted)

**Question 2** Given the following:

$$\widehat{sat} = 1,028.10 + \frac{19.30}{6.29}hsiz e - \frac{2.19}{0.53}hsiz e^2 - \frac{45.09}{4.29}female - \frac{169.81}{12.71}black \\ + \frac{62.31}{18.15}female * black$$

$$n = 4,137 \quad R^2 = 0.0858$$

The variables are defined as:

*sat* = Combine SAT score.

*hsiz e* = high school graduating class size in the hundreds.

*female* = Whether the student is female or not.

*black* = Whether a student is black or not.

1. Is there strong evidence that  $hsiz e^2$  should be included in the model? From this equation, what is the optimal high school size?
2. Holding *hsiz e* fixed, what is the estimated difference in SAT score between *non-black* females and *nonblack* males? How statistically significant is this estimated difference?
3. What is the estimated difference in SAT score between nonblack males and black males? Test the null hypothesis that there is no difference between their scores, against the alternative that there is a difference.
4. What is the estimated difference in SAT score between black females and nonblack females? What would you need to do to test whether the difference is statistically significant?

## Computer Exercises

**Question 3** Use the data in WAGE2 for this exercise.

Given the following:

$$\begin{aligned} \log(wage) = & \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure \\ & + \beta_4 married + \beta_5 black + \beta_6 south + \beta_7 urban + u \end{aligned}$$

1. Report the estimated model. Holding other factors fixed, what is the approximate difference in monthly salary between blacks and nonblacks? Is this difference statistically significant?
2. Add the variables  $exper^2$  and  $tenure^2$  to the equation and show that they are jointly insignificant at even the 20% level.
3. Extend the original model to allow the return to education to depend on race and test whether the return to education does depend on race.
4. Again, start with the original model, but now allow wages to differ across four groups of people: married and black, *married* and *nonblack*, *single* and *black*, and *single* and *nonblack*. What is the estimated wage differential between *married blacks* and *married nonblacks*?