503/650 2016: Assignment 11 – CRM Points missed:

**Your name: Name of TA:**

* **DUE in Scott Long’s mailbox in BH 744 by noon on Tuesday of exam week.**
* Put your assignment in a campus mail envelope **addressed to YOUR campus address** (there are envelopes available in the Sociology office *under* the faculty mailboxes in BH 744).

1. \_\_\_ of 10: Use the variables for the CRM that you selected in Assignment 3. Count variables such as # of children that have a limited range, variables with a spike at one of the values, or variables with the majority of cases at large integers often cause problems and should not be used. The range of your count variable must be 7 or greater. One independent variable must be a factor variable F; one must be continuous variable C. X is a vector of 3 to 5 variables that can be either continuous or categorical. Include a publishable table describing your variables, including the distribution of the dependent variable (show percentages not raw numbers). If your outcome has many values, you can group counts for higher values in your table (e.g., show 0, 1, 2, 3, 4, 5-9, 10-15, etc. in the rows) or use a graph. Your analysis would not use the grouped counts. If your outcome has extreme values (e.g., a few 99s), consider top coding these observations.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 5: Use poisson and nbreg to regress Y on C, F, and X. C and F must be statistically significant in the NBRM. Compute the standardized and unstandardized coefficients for F and C for the NBRM using listcoef, help. Include the output from listcoef here.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 15: For the NBRM, interpret the standardized factor change coefficient for C and the unstandardized factor change coefficients for C and F. Include statistical significance. Each sentence should read as if it were part of a published article. You would interpret the coefficients for the PRM in the same way, but only interpret the NBRM coefficients for this question.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 10: For the NBRM, use mchange for F and C using AME. Interpret the average marginal effects for F as though it were part of a paper. Interpret one of the average discrete changes for C as though it were part of a paper. Include statistical significance.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 5: Which method of interpretation, factor change or marginal effects, do you find most effective? Briefly explain why. Discuss how your results differ, even subtly, between the two.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 5: Test the NBRM against the alternative of the PRM. Write the result as though it were part of a research paper. This can probably be accomplished in a single sentence.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 10: Plotting probabilities.

**7a.** Use mgen to create a plot comparing Pr(0) from the PRM and the NBRM. Show the graph here. See the notes or Long and Freese for the commands you can use.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

**7b.** Use mgen, meanpred to create a plot comparing fits of NBRM and PRM. Show the graph here.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

**7c.** Describe the substantive implications of what you find.

Font for your answer

1. \_\_\_ of 10: Comparing models

**8a.** Use countfit to compare PRM, NBRM, ZIP and ZINB. Include the output and graph from countfit. Consider using the maxcount() option if your dependent variable has many more than the default of 9 counts.

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

**8b.** Based on these results and your substantive understanding of the process being studied, which model would you use and why?

Font for your answer

Paste any Stata output here in Courier New 9 point font, single-spaced

1. \_\_\_ of 10: Assessment of the overall effectiveness of your answers.
2. \_\_\_ of 20: Overall assessment of your research log and file structure on Box. You need to make certain that all of your work is up to date, organized, and placed on the Box. You must include a folder with all of the do-files and log files for all assignments. These must run without error with no changes made to the files.

Source: CDAiu2016-a11-CRM-2016-07-26.docx