

Soc 73994, Homework #3

Interpreting results: Adjusted Predictions and Marginal effects

Richard Williams, University of Notre Dame, <https://www3.nd.edu/~rwilliam/>

Last revised September 10, 2024

All answers should be typed and submitted through Canvas. Be sure your response includes your name, the date, and a clear title, e.g. Homework # 3. If there is a huge amount of output for any analyses you run yourself, you may want to be selective in what you copy and paste into your assignment (but make sure you include enough so it is clear what commands you executed, e.g. you might show all the commands but only parts of the output).

This assignment has two parts. First, you will answer questions about a hypothetical logistic regression analysis. We will again use HW # 2's (fake) data on the Ted Cruz for President in 2032 campaign. You will then conduct and interpret a similar analysis using the data set of your choice. If you don't understand what a command is doing, check the help file and/or the class handouts. You will need to run the commands yourself since the output is not included here.

1. The Cruz campaign demands only the highest quality research. It is therefore spending millions of dollars to have Sociology Graduate Students write Stata programs for it. Unbeknownst to the campaign, a still-bitter Donald Trump Jr. supporter has infiltrated the research team and is deliberately trying to sabotage the effort.

a. Different graduate students have proposed the following ways for coding the logistic regression analysis. Eventually the `margins` command will be used to gain additional insights. The Trump supporter is arguing that it does not make any difference which coding you use because they all give the same results and therefore you might as well use one of the first two methods, which she coded. Explain why her advice is wrong.

```
use https://www3.nd.edu/~rwilliam/statafiles/cruz, clear
```

```
* 1A
```

```
* Approach 1 - Generate dummies yourself
```

```
tab ses, gen(ses)
```

```
logit cruz male socialcons fiscalcons ses2 ses3 , nolog
```

```
* Approach 2 - Let xi: generate the dummies
```

```
xi: logit cruz i.male i.socialcons fiscalcons i.ses, nolog
```

```
* Approach 3 - Yet another approach that somebody insists on trying
```

```
logit cruz i.male i.socialcons fiscalcons i.ses , nolog
```

b. The Trump supporter is outvoted and the third method is used. The following `margins` commands are then suggested to make the effects of the continuous variable `fiscalcons` easier to interpret. The Trump supporter says the first `margins + marginsplot` commands are way too complicated and they should just use her 2nd `margins` command instead. Explain what each of these approaches does and which you think is better.

```
* 1B
* Possible margins & marginsplot commands.
use https://www3.nd.edu/~rwilliam/statafiles/cruz, clear
logit cruz i.male i.socialcons fiscalcons i.ses , nolog
margins, at(fiscalcons = (-12(3)12))
marginsplot, noci name(g1)
margins fiscalcons
```

c. The campaign has gotten a little suspicious of the undercover Trump Jr. supporter and has therefore reassigned her to do door to door campaigning in Barrow, Alaska, which happens to be one of the coldest cities in the world. With her gone, the campaign runs three sets of analyses. Briefly explain what each set of analyses is doing and interpret the results.

```
* 1C
* 1. Basic margins commands
use https://www3.nd.edu/~rwilliam/statafiles/cruz, clear
logit cruz i.male i.socialcons fiscalcons i.ses , nolog
margins male socialcons ses
margins, dydx(*)

* 2. Adjusted predictions for male social conservative fiscal moderates
* at different SES levels
use https://www3.nd.edu/~rwilliam/statafiles/cruz, clear
logit cruz i.male i.socialcons fiscalcons i.ses , nolog
quietly mtable, at(male=1 socialcons=1 ses=1 fiscalcons=0) rowname(Low SES)
quietly mtable, at(male=1 socialcons=1 ses=2 fiscalcons=0) below rowname(Med SES)
mtable, at(male=1 socialcons=1 ses=3 fiscalcons=0) below rowname(High SES)

* 3. Examine effects of fiscalcons both by itself and with gender
use https://www3.nd.edu/~rwilliam/statafiles/cruz, clear
logit cruz i.male i.socialcons fiscalcons i.ses , nolog
mcp fiscalcons, var1(20) plotopts(name(g2, replace))
mcp fiscalcons male, var1(20) plotopts(name(g3, replace))
```

2. Now you will do similar analyses using a data set of your choice (preferably the data you want to use for your paper but you can pick anything, including the data you used in HW 1 or 2). From this data set you will need

- A binary dependent variable. If you are desperate, remember that continuous, ordinal, and count variables can be dichotomized if necessary, e.g. an ordinal variable might be recoded to 1 = agree, 0 = disagree.
- Three or more independent variables. At least one should be categorical (but it can have more than 2 categories, e.g. you could use religion if it doesn't have too many categories) and at least 1 should be continuous (or if desperate, use an ordinal variable you treat as continuous). You can have more than three variables if you want.
- Briefly explain what each of these variables are and how they are coded. You can copy and paste from your earlier assignments if you want.

Now do the following:

- a. Run a logistic regression model. Make sure you use factor variable notation correctly.
- b. Estimate the adjusted predictions for your categorical variables. Also estimate the marginal effects for all your variables. For your categorical variables explain how the adjusted predictions and the marginal effects are related. Also briefly justify your decision to use (or not use) the atmeans option.

- c. Compute 3 or more adjusted predictions at representative values. For example, you could compute APRs for a male Black person, a female Black person, a male White person, and a female White person. You can use the `margins` command for this or (possibly better) Long and Freese's `mtable` command.
- d. Use the `mcp` command (or else `margins` and `marginsplot`) to plot the adjusted predictions at different values for your continuous variable. If you think it is desirable, you can expand or narrow the range for your predictions (e.g. exclude extreme outliers, or go beyond the range of the observed values if you think it is useful to do so). But explain your reasoning if you do so.
- e. Again use the `mcp` command (or else `margins` and `marginsplot`) to plot the adjusted predictions for your continuous variable separately for the different values of one of your categorical variables, e.g. do separate lines for men and women.
- f. (Optional) Do any other analyses you think might be helpful using the `margins` or `marginsplot` or `mcp` commands or one of Long and Freese's commands. For example, you might try out their `mchange` command. Interpret your results.