Advanced Topics in Statistical Methods

PLSC 504

Exercise One
September 11, 2017

Preliminaries and Data

The time was January, 2005. Condoleeza Rice was sworn in as the first African-American Secretary of State, Mahmoud Abbas was declared the winner of the Palestinian election, and (perhaps most important) Texas light sweet crude was selling for the princely sum of $45 a barrel. It was during those innocent, happier days that ABC News and the Washington Post commissioned a poll about public opinion on traffic. Among other things, pollsters asked 1204 lucky, randomly-selected Americans:

> "What kind of vehicle do you usually drive – a car, an SUV, a pickup truck, or what?"

What does this have to do with political science? The answer ought to be obvious.[1] We'll explore the political dynamics of car ownership, using the data from the 2005 ABC/WP poll. The main variable of interest is `cartype`, coded one for cars, two for SUVs, and three for pickup trucks. Covariates include dummy variables for `urban` residence, being `married`, having `kids`, and being `black` and/or `female`, as well as a naturally coded variable for `age` and an ordinal variable for level of `education`. Best of all, we also have two dichotomous variables for political party (`democrat` and `GOP`, with independents as our baseline) and a four-point ordinal scale indicating each respondent's approval or disapproval for then-President Bush.

Exercise

1. After examining summary statistics, start with a (probably untenable) assumption: That the ordering of values for `cartype` (with car = 1, SUVs = 2, and pickups = 3) is reasonably thought of as ordinal. Fit and interpret a model of vehicle ownership that reflects this assumption.

2. Next, estimate a multinomial logit (MNL) model of vehicle type ownership. Report your estimation results, and interpret these findings, in statistical and substantive terms. Are the results in the "expected directions"? Discuss their statistical significance.

3. Using the MNL results, generate and examine the predicted probabilities of each type of vehicle, across the range of values for some of your more important independent variables, using tables or graphs of the probabilities. Interpret these results in substantive terms.

4. Rerun the MNL model, changing the baseline category, and discuss your results. How (if at all) does the interpretation change? In your opinion, does respecifying the baseline category tell you anything you didn't know before, or make anything clearer?

5. Test for whether the data/model conform to the MNL model's IIA assumption. Use whatever tests you can / are aware of, and discuss your findings on this point in both statistical and substantive terms.

6. Irrespective of the results of the IIA test(s), reestimate the same specification using a multinomial probit model. Compare those findings to the MNL results, and briefly discuss similarities and differences.

7. Finally, briefly discuss *in substantive terms* what your statistical conclusions suggest for the relationship between political ideology/preferences and automobile ownership.

This exercise is due by 5:00 p.m. ET on Tuesday, September 18, 2017, and is worth the usual 50 points.

[1] Example: A Google search for `"suv-driving republican -democrat"` yields 71,200 hits. For `"suv-driving democrat -republican"`? 20,800 hits. (Similar results obtain for pickups...).